(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US];

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP) 3 February 2000 (03.02.2000) Filed on

US Filed on

09/560,875 (CIP) 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom -{US/US}; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

Glovsky, and Popeo, P.C., One Financial Center, Boston,

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris,

MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,

LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,

NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,

TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

Published:

without international search report and to be republished upon receipt of that report

CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

5

10

15

20

25

30

35

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers. than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that. have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

10

15

20

25

30

35

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

5

10

15

20

25

30

35

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5

10

15

20

25

30

35

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

25 4.3 ANTISENSE

5

10

15

20

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

5

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

5

10

. 15

20

25

30

35

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

5

10

15

20

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

5

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

25

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

35

10

5

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer 20 programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. 25 Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available 30 from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

5

10

15

20

25

30

35

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

5

10

15

20

25

30

35

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

5

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

- J,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

10

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No.

WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

30

35

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

10

15

20

25

30

35

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeck, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

10

15

20

25

30

35

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

5

10

15

20

25

30

35

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (*i.e.*, increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

5

10

15

20

25

30

35

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

10

15

20

25

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particularneurotoxins; and
 - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

15 4.10.19 IDENTIFICATION OF POLYMORPHISMS

10

20

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

5

10

15

20

25

30

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

10

15

20

25

30

35

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

5

10

15

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

5

10

15

20

25

30

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

10

15

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-\alpha and TGF-\beta), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

5

10

15

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1 , IgG_2 , and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation, See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, 35 · synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol., 133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

5

10

15

20

25

30

35

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al., (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragment's and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_{v} fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced antitumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

5

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

5

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
 - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

5

10

15

20

25

30

35

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategics. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer et al. (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ($CviII^{**}$), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a $CviII^{**}$ digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that $CviII^{**}$ restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

· 10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

20 5.0 EXAMPLES

5

10

15

25

30

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virgimia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
	1		976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
			217 225 238 271 317 404 446 469 503
			513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
			147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
		,	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
		1	658 666 683 725 742 764 767 786 801
		1	805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
	1		1128 1142 1162 1181-1192 1199 1204
	1	ł	1218-1219 1225 1232 1253 1267 1271-
		<u> </u>	1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
			566 596 663 670 746 798 816-819 876
			892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes		<u> </u>	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
			240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
		·	1003 1067-1070 1118 1156 1193-1200
3 3. 1	<u> </u>	1 777 000	1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
			118 129 132 138 151 158-163 182 195-
	,		203 215 217 238 264 269 353 384 398
			408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
			671-672 722 734 757-773 815 828-835
	ŀ		874 891 898 919 926-927 976 988
		1	1021 1037 1041 1062 1067 1071 1080
		1.	1083 1093 1122 1131 1185 1201 1254
adala lai da	CIDCO	AVDOOL	1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
	1		107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
	}		446 454 477 504-505 509 514 518-519
	1		535 537 564 574-583 620-627 639 653
	1		673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
achilt kid-as	Invita	AKTOOS	1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
			446 487 564 575 844 868 910 927 976
achilt lung	CIRCO	AT COOL	1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514

580 582 592 594 634 -678 725 851 873 918 67 1076 1083 1152 80-182 188 215 537 -682 789 804-810 868 1042 1059 1335 215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 6722 735-744 761 771 6868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120 1174 1224 1268 1331
67 1076 1083 1152 80-182 188 215 537 -682 789 804-810 868 1042 1059 1335 215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 1537 549-550 564 566 618 638 657 667 681 1722 735-744 761 771 1868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
30-182 188 215 537 -682 789 804-810 868 1042 1059 1335 215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 1537 549-550 564 566 618 638 657 667 681 1722 735-744 761 771 1868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
-682 789 804-810 868 1042 1059 1335 215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 1537 549-550 564 566 618 638 657 667 681 1722 735-744 761 771 1868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
1042 1059 1335 215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 1537 549-550 564 566 618 638 657 667 681 1722 735-744 761 771 1868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
215-216 238 446 514 683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 6722 735-744 761 771 6868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
683-684 698 753 798 927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
927 976 1038-1039 1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
1245 1256 05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
05 384 468-469 496 753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
753 832 844 941-942 256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
256 1293 51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
51 71 74 79-80 101 -125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
-125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
-125 138 140 143-149 -212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
-212 215-217 238 264 445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
445-446 496 504 512 537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
537 549-550 564 566 618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
618 638 657 667 681 722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
722 735-744 761 771 868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
868 875-876 918 926- 976 1023 1042 1048 1076 1083 1117 1120
976 1023 1042 1048 1076 1083 1117 1120
1076 1083 1117 1120
641 700
140 151 185 217 238
504 514 534 545 549
952 976 1041-1042
1152 1224
26 140 151 183 215
642 701-706 811 877
3 1117 1131
405 409 414 496 545
952 1178 1329-1335
79 108 111 116 137
213-215 238 305-307
466 516 519 534 538-
-554 566 584 586 592
-629 643-645 652 707-
866-871 873 919 927
1034 1042 1064 1083
1152 1225 1229 1268
77 105 111 120 122
77 105 111 129 132 545 549 581 598 628
844 860 868 873 919
976 1042 1111 1141
1266 1346
1083
194 536 545 564 592
976 1042 1152 1268
710 1044 1134 1400
51 205 207 238 332-
-401 440 466 470-471
832 877 927 976 1006
1134 1192 1202-1205
1104 1172 1202-1203
i
; 173Z 844 0 9 3 4 3 T 1 9 3 7 7 4 2 5 2 9

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
	1 22.000	221001	138 151 204-206 215-217 238 269 316
]]	414 433 505 510 513 550 555 580 582
ĺ	1	1	596 675 722 745 798 814 836-841 851
1			918 976 1041 1043 1073 1083 1131
			1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic	}	
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		
of chromosome 8	Research		<u> </u>
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic]	
of chromosome 8	Research	FDD 400 5	
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic		1
of chromosome 8	Research BioChain	ESCONO	74 128 228
esophagus fetal brain		ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004 FBR006	215 893 927 1001 48 61 101 120 132 138 140 147 208
Terat night	Cionicci	סטטאמז	225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
1			829 836 859 909 927 943 947 963 1057
			1067-1068 1104 1135-1140 1162 1206-
1	}		1207 1235 1268 1288 1307-1308 1319
			1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
			535 683 761 798 820-827 844 876 909
			963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
Cotal lines 1	Cobout	ET COOL	859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
	University		69-71 74 77 79 87-90 101 107 110-111
	Į l		114 120 128-131 138 140 147 150-155 197 210 215 217 225 238 312 367 384
	1		414 440 446 460 468 483 496 504-507
	1		511-515 518-519 523 533-535 537 541
	}		544-545 547-550 555-560 564 566 571
	1		577 582 585-586 598 636 646-647 649
			652 664 698 709-710 714 722-723 731
	[735-736 746-753 761 784 798 823 829
	1		832 844 851 858-859 868 873 876 898
	Ì	•	927 943 949 952 963 976 984 1002
			1021 1023 1040 1042 1044 1050 1083
			1093 1116 1120 1129 1131 1144 1174
	1		1217 1251 1254 1256 1302 1308 1311
	<u> </u>		1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
	University		111 120 129 147 207 210 215-216 238
			250 330 353 359 366 383-384 414 478
			505 508-509 511 515-524 534-535 537 544-545 564 566 571 577 591 598 638
	, .		

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798
			851 859 873 876 909 927 949 952 983-
	1		984 1002 1023 1042-1044 1085 1095
	1		1131 1144 1178 1199 1233 1240-1270
			1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
			580 722 730 749 844 918 943 976 1051
			1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141
			1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
	1		225 264 316 405 422-429 488-494 496
	1		519 534-535 537 566 675 732 859 876-
	1	•	877 898 947 949-950 963 976 1001 1062 1076 1083 1117 1144 1165 1268
	j		1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
ambinoar cora	DioCham	1 00001	316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
l			976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
			138 145 151 188 197 207 215 238 264
	i	1	271 294 316 367 414 440 446 466 504
			513-514 535 542-543 550 564 571 596
			635 648-654 675 711-715 722-723 798
			832 872 876 883 927 976 1095 1144
macrophage	Invitrogen	HMP001	1168 1171 1178 1211 1335
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
miani Olani	University	102002	151 161 175-179 185 216-217 264 295
	Oniversity		299 308-310 371-373 462 476 504 511-
			513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
			858-859 876 898 909 949 976 1045-
			1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
_	<u> </u>		1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
			1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
	University	177.004	
infant brain	Columbia University	IBS001	51 111 376 474 790 876 949 1144 1204
lung fibroblest		I EDOO!	1221
lung, fibroblast	Strategene Invitrogen	LFB001 LGT002	151 316 462 514 534 582 675 939 1131 1-7 41 74 79 94 115 120 138-139 156
		LOTOG .	
lung tumor	HIVINOSCH	j.	1 715 717 760 700 700 202 227 274 275 204 1
lung tumor	Mythogen		215 217 269 280 296 337 374-375 384
rung tumor	mvinogen		404 446 454 475-480 498 514 518-519
ning tumor	III VIII OGEII		404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705
ning tumor	nvadogen		404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705 721-724 754-756 779 859 868 872-874
lung tumor	IIVIIIOgen		404 446 454 475-480 498 514 518-519 522 537 545 564 577 597 653 658 705

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
j			634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
			147 151 212 215 218 238 252 288 312-
		}	314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
			564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
	1		836 841 859 866 873-874 882-883 918-
			919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
			657 660 729 873 883 927 952 963 1008
			1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL			919 929 939 952 976 1071 1118 1218
1424			1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
			217 250-256 264 297-299 305 377-378
	ļ		398 446 481-486 505 512 537 545 549
	i		571 592 725 730-733 816 829 836 844
	·		868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
			1055 1076 1083 1091 1093 1116-1117
			1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
	Dudiogonio	11,200,	1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells	Jumegene		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
			1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
			1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
, <u>G</u>			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
			963 967-978 1042 1120 1152 1223-
•			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
opinar cora	Cionicon	57 0001	270 343-344 353 379 516 537 566 740
			828 927 976 979-994 1092 1153-1159
			1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
этопроп	Clonicon	310001	995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
ermaning	Ciontecn	1117002	963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
uryinus .	Cionetecu	וממומנו	750 867 874 878-881 927 963 1023
thumus	Clentach	TUMOO	1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306 317-319 336 340 359 380 398 446 448-
,			463 512 519 545 554 587 598 698 724-
			I I
	L	<u>. L </u>	725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
			1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
	j		210 217 222 253 264 271 277-286 294
		1	320-326 345-352 361 381-382 446 467
			483 514 534 549-550 564 578 602 649
			844 882-883 927 950 956 976 1008-
			1028 1076 1083 1117-1120 1142 1163-
		<u> </u>	1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
			952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
			885-886 976 1001 1032-1033
			1232

TABLE 2

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen	293	76
			(CDR62) encoded by clone pY2.		
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein	293	100
			kinase (ortholog of mouse and rat MAK (male		1
			germ cell-associated kinase))		
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus	prostaglandin F2a receptor regulatory protein	569	89
		norvegicus	precursor		1
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

D No.	SEQ	Accession	Species	Description	Smith-	1%
NO:			Species	Description		Identity
197 G04067 Homo sapiens Human secreted protein, SEQ ID NO: 8148, 83 42		140.	1			lucinity
Human secreted protein, SEQ ID NO: 6953. 116 72 73 766688 Homo supiens Human secreted protein, SEQ ID NO: 7452. 96 67 73 766688 Homo supiens Human secreted protein, SEQ ID NO: 7303. 58 32 766688 Homo supiens Human secreted protein, SEQ ID NO: 7303. 58 32 766688 Homo supiens Human secreted protein Requence SEQ ID NO: 130 348 95 766688 Homo supiens Human secreted protein sequence SEQ ID NO: 130 148 95 76 77 147 77 77 77 77 77 7		G04067	Homo capiens	Human secreted protein SEO ID NO: 8148		42
131 G03371 Homo sapiens Human secreted protein, SEO ID NO: 7452. 96 67						
192 193 194 195			Homo sapiens			1
Y87071						
NO:110.						
35	5.	10,0,,	Tromo Saprons		15.0	1 "
136	35	1)15131	Homo saniens		182	48
Sequence SEQ ID NO:150. Sequence SEQ ID NO:150. Sequence SEQ ID NO:150. Sequence SEQ ID NO:160. Sema domain, immunoglobulin domain (lg), transmembrane domain (TM) and short typolasmic domain) Sequence SEQ ID NO:160. Semantic domain) Sequence SEQ ID NO:160. Sequence SEQ ID NO:1709. Sequence Sequenc						
AL	30	17,5401	Tromo suprems		702	
	37	AL 133215	Homo saniens		687	99
	٥,	1.2.052.0	Tronso supremb		33.	
AC067969 amino acids 3338-4088 Homo sapiens ryanodine receptor I (skeletal) 386 66 66 3338-4088 Homo sapiens Homo sapiens		1	1			
338-4088	38	AC067969	amino acids		386	66
ALO31588	00	1.1000.707			1000	1
FGENES and GENEWISE	39	AL031588	- L	dJ1163J1.1 (mostly supported by GENSCAN.	493	76
40 G03628 Homo sapiens Human secreted protein, SEQ ID NO: 7709. 110 51			1			
AF132969 Homo sapiens CG1-35 protein 228 68 43 X61048 Hydra sp. Human secreted protein encoded by gene 45. 220 88 mini-collagen 105 35 35 44 M76546 Helianthus annuus hydroxyproline-rich protein 110 31 31 31 31 32 33 34 34 34 34 34 34	40	G03628	Homo sapiens		110	51
42 Y36268						68
AS X61048						88
Main						
A						
45				- January Francisco	1	
Selegans	45	U82288	Caenorhabditi	Rac-like GTPase	139	70
A6 G03477 Homo sapiens PRO0657 Homo sapiens PRO0657 Homo sapiens PRO0657 Homo sapiens PRO0657 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 90 59		1				
AF090942	46	G03477		Human secreted protein, SEQ ID NO: 7558.	118	58
AJ005560 Mus musculus SPR2B protein 72 56	47	AF090942			113	63
AJ005560 Mus musculus SPR2B protein 72 56	48		Homo sapiens	Human secreted protein, SEO ID NO: 7645.	90	59
Social Color Homo sapiens Human secreted protein, SEQ ID NO: 6531. 385 98						56
Signature Sign			musculus	•		
51 Y91649 Homo sapiens Human secreted protein sequence encoded by gene 60 SEQ ID NO:322. 973 94 52 U93563 Homo sapiens putative p150 105 38 53 Y55927 Homo sapiens Human STLK2 protein. 699 85 54 G02607 Homo sapiens Human STLK2 protein. 699 85 55 AB008175 Mus musculus Human secreted protein, SEQ ID NO: 6688. 145 56 56 M68941 Homo sapiens protein-tyrosine phophatase 165 41 57 AL031600 Homo sapiens c390E6.1 (chloride channel 7) 338 76 58 AF011417 Mus musculus putative pheromone receptor 143 55 59 AF167320 Mus musculus interferon regultory factor 7 263 96 60 U73036 Homo sapiens Human secreted protein clone cb98_4. 791 98 62 Y29861 Homo sapiens Human secreted protein clone cb98_4. 791 98 6	50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
Seminorm	51	Y91649			973	94
52 U93563 Homo sapiens putative p150 105 38 53 Y55927 Homo sapiens Human STLK2 protein. 699 85 54 G02607 Homo sapiens Human secreted protein, SEQ ID NO: 6688. 145 56 55 AB008175 Mus hepatic nuclear factor 1-beta short form 356 74 56 M68941 Homo sapiens protein-tyrosine phophatase 165 41 57 AL031600 Homo sapiens c390E6.1 (chloride channel 7) 338 76 58 AF011417 Mus putative pheromone receptor 143 55 59 AF167320 Mus zinc finger protein ZFP113 558 68 60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98 4. 791 98 63 U35376 Homo sapiens	-					
53 Y55927 Homo sapiens Human STLK2 protein. 699 85 54 G02607 Homo sapiens Human secreted protein, SEQ ID NO: 6688. 145 56 55 AB008175 Mus musculus hepatic nuclear factor 1-beta short form musculus 356 74 56 M68941 Homo sapiens protein-tyrosine phophatase 165 41 57 AL031600 Homo sapiens c390E6.1 (chloride channel 7) 338 76 58 AF011417 Mus musculus putative pheromone receptor 143 55 59 AF167320 Mus musculus zinc finger protein ZFP113 558 68 60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98_4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapie	52	U93563	Homo sapiens		105	38
Section	53	Y55927	Homo sapiens		699	85
AB008175 Mus musculus hepatic nuclear factor 1-beta short form 356 74	54	G02607			145	56
M68941 Homo sapiens Protein-tyrosine phophatase 165 41	55	AB008175		hepatic nuclear factor 1-beta short form	356	74
57 AL031600 Homo sapiens c390E6.1 (chloride channel 7) 338 76 58 AF011417 Mus musculus putative pheromone receptor 143 55 59 AF167320 Mus musculus zinc finger protein ZFP113 558 68 60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98 4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens REB2 614 100 68 AF030027 <td></td> <td></td> <td>musculus</td> <td></td> <td></td> <td>1</td>			musculus			1
57 AL031600 Homo sapiens c390E6.1 (chloride channel 7) 338 76 58 AF011417 Mus musculus putative pheromone receptor 143 55 59 AF167320 Mus musculus zinc finger protein ZFP113 558 68 60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98 4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95	56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
58 AF011417 Mus musculus putative pheromone receptor 143 55 59 AF167320 Mus musculus zinc finger protein ZFP113 558 68 60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98 4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta sexta 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB014885	57	AL031600	Homo sapiens		338	76
Mus	58	AF011417		putative pheromone receptor	143	55
musculus musculus			musculus	· · · · ·		
60 U73036 Homo sapiens interferon regultory factor 7 263 96 61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98_4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP sexta 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 </td <td>59</td> <td>AF167320</td> <td>Mus</td> <td>zinc finger protein ZFP113</td> <td>558</td> <td>68</td>	59	AF167320	Mus	zinc finger protein ZFP113	558	68
61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98_4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454			musculus			
61 X07984 Mus musculus protein-tyrosine kinase 297 69 62 Y29861 Homo sapiens Human secreted protein clone cb98_4. 791 98 63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Ca	60	U73036	Homo sapiens	interferon regultory factor 7	263	
musculus	61	X07984	Mus	protein-tyrosine kinase	297	69
63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi phospholipase B 955 73 73 AF045454 Cavia porcellus phospholipase B 955 73						
63 U35376 Homo sapiens repressor transcriptional factor 485 65 64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi phospholipase B 955 73 73 AF045454 Cavia porcellus phospholipase B 955 73			Homo sapiens	Human secreted protein clone cb98_4.		
64 AF265555 Homo sapiens ubiquitin-conjugating BIR-domain enzyme APOLLON 785 74 65 G03883 Homo sapiens Human secreted protein, SEQ ID NO: 7964. 88 95 66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73	63	U35376	Homo sapiens	repressor transcriptional factor	485	65
APOLLON			Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	785	74
66 AF177390 Manduca sexta antennal specific membrane protein AMP 274 54 67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesyirus 4 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73						
67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73					88	95
67 AB040800 Homo sapiens SREB2 614 100 68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73	66	AF177390	Manduca	antennal specific membrane protein AMP	274	54
68 AF030027 Equine herpesvirus 4 24 213 26 69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73						<u>L</u>
herpesvirus 4		1			_ 1	100
69 G02965 Homo sapiens Human secreted protein, SEQ ID NO: 7046. 261 95 70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73	68	AF030027		24	213	26
70 W75770 Homo sapiens Human oxidoreductase YTFO3. 1144 98 71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73						L
71 AB011135 Homo sapiens KIAA0563 protein 239 76 72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73						
72 AB014885 Halocynthia roretzi HrPOPK-1 813 78 73 AF045454 Cavia porcellus phospholipase B 955 73		1	Homo sapiens	Human oxidoreductase YTFO3.		
73 AF045454 Cavia phospholipase B 955 73 porcellus		AB011135	Homo sapiens		239	76
73 AF045454 Cavia phospholipase B 955 73 porcellus		AB014885	Halocynthia	НгРОРК-1	813	78
porcellus		L	roretzi			
	73	AF045454	Cavia	phospholipase B	955	73
74 J02870 Mus Iaminin receptor . 308 61		<u> </u>	<u> </u>	·	1	
	74	J02870	Mus	laminin receptor	308	61

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	l			Score	
	<u> </u>	musculus			<u> </u>
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
	15110054	norvegicus			-
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
	1 1/20422	77	complex component TRAP240	460	l
77	Y38422 Y14596	Homo sapiens	Human secreted protein. Human T-type voltage-gated Ca channel alpha-	1357	76
78	1 14396	Homo sapiens	1-I (hCavT3).	1337	99
79	Y14591	Human	APM-1 protein	767	100
	11.551	papillomaviru	14 M 1 process] '`.	
		s type 68		1	j
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
		thaliana			i
82	L46815	Mus	DNA binding protein Rc	895	75
		musculus			j
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
		<u> </u>	designated HSCOP-6.	<u> </u>	
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
-	1		sequence SEQ ID NO:100.	4200	<u> </u>
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
00	AF177203	musculus	HAC3	200	-
89 90	Y28280	Homo sapiens Homo sapiens	cerebral cell adhesion molecule Human G-protein coupled receptor GRIR-2.	290 326	56 79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
	7.13232	brucei	GIT phosphonpase e (FBTT 350)	1	1 "
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
		norvegicus		ì	
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	675	48
		norvegicus	kinase		<u> </u>
100	AF279265	Homo sapiens	putative anion transporter I	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	60
100	U22829	1,	difference at residue 58		ļ.,,
102	022829	Mus musculus	P2Y purinoceptor	264	42
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
105	143023	Homo sapicus	receptor-B3.	310	"
				707	1 00-
104	Y94990	Homo saniens	l Human secreted protein vb21 1. SEO ID NO 20	1 /8/	198
104	Y94990 Y87342	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20. Human signal peptide containing protein HSPP-	787 343	98
104 105	Y94990 Y87342	Homo sapiens Homo sapiens	Human secreted protein Vb21_1, SEQ ID NO:20. Human signal peptide containing protein HSPP- 119 SEQ ID NO:119.	343	57
			Human signal peptide containing protein HSPP-		
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119.	343	57
105 106	Y87342 AF169312	Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein	343	57 67
105 106 107 108	Y87342 AF169312 AF116657 AE000401	Homo sapiens Homo sapiens Homo sapiens Escherichia coli	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter	343 212 74 587	57 67 52 96
105 106 107 108	Y87342 AF169312 AF116657 AE000401 Y38395	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10.	343 212 74 587	57 67 52 96
105 106 107 108	Y87342 AF169312 AF116657 AE000401	Homo sapiens Homo sapiens Homo sapiens Escherichia coli	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone	343 212 74 587	57 67 52 96
105 106 107 108 109 110	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence.	343 212 74 587 693 182	57 67 52 96 100 94
105 106 107 108 109 110	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153	343 212 74 587 693 182	57 67 52 96 100 94
105 106 107 108 109 110	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein	343 212 74 587 693 182	57 67 52 96 100 94
105 106 107 108 109 110 111 112	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84.	343 212 74 587 693 182 464 274	57 67 52 96 100 94 85 51
105 106 107 108 109 110 111 112	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td	343 212 74 587 693 182 464 274	57 67 52 96 100 94 85 51
105 106 107 108 109 110 111 112	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown	343 212 74 587 693 182 464 274 301 520	57 67 52 96 100 94 85 51 71 75
105 106 107 108 109 110 111 112 113 114 115	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956 M83738	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown protein-tyrosine phosphatase	343 212 74 587 693 182 464 274 301 520 251	57 67 52 96 100 94 85 51 71 75 92
105 106 107 108 109 110 111 112	Y87342 AF169312 AF116657 AE000401 Y38395 Y78801 Z25535 Y94939 AF016365 AC007956	Homo sapiens Homo sapiens Homo sapiens Escherichia coli Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119. hepatic angiopoietin-related protein PRO1310 sialic acid transporter Human secreted protein encoded by gene No. 10. Hydrophobic domain containing protein clone HP00631 amino acid sequence. nuclear pore complex protein hnup153 Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84. hexokinase 1 isoform td unknown	343 212 74 587 693 182 464 274 301 520	57 67 52 96 100 94 85 51 71 75

OFO.	1 Assession	Species	I Daniel Barrell	1 C '4L	1%
SEQ ID	Accession No.	Species	Description	Smith-	
NO:	NO.			Waterman	Identity
	1,4,0,0	·		Score	
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
		norvegicus		<u> </u>	1
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr	1646	94
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi s elegans	contains similarity to C2 domains	219	29
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit 4	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IF116b	496	67
131	AF201734	Mus	testis specific serine kinase-3	800	87
		musculus			
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1157	87
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860)	5041	99
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDI1-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	392	61
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
160	AJ250425	Rattus norvegicus	Collybistin I	556	74
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100
	1 002003	1 TOTALO SAPIETIS	Trainian scarcing protein, SEQ ID NO. 0300.	7,0	1,00

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:	j.			Score	
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus norvegicus	calcium transporter CaT1	700	96
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus	NIK-related kinase	197	43
169	AF252293	musculus	PAR3	596	44
170	U59429	Homo sapiens		481	82
		Cricetinae gen. sp.	diacylglycerol kinase eta		
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55 .
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
			codon)	175	80
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor		
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit	157	72
193	M92084	Theileria	polypeptide (MSP)GPIIb-IIIa. casein kinase II alpha subunit	364	50
104	VICEAS	parva	Mambana haund	448	90
194 195	Y66645 W95631	Homo sapiens Homo sapiens	Membrane-bound protein PRO1310. Homo sapiens secreted protein gene clone	382	49
196	AF255614	Rattus	hj968_2. scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
	1 17000000	thaliana		1316	12
198	AF073967	Mus musculus	olfactory receptor	316	43
	<u> </u>	domesticus		<u> </u>	
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1		Waterman	Identity
NO:		į		Score	1
			{ovarian cancer critical region of deletion}		1
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
210	D19992	rionio sapiens		341	02
011	A D117040	177	protein, calphotin.	1348	99
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C		1 * *
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
	<u> </u>	norvegicus	neurofascin		
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	933	93
		musculus			
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus	prostaglandin F2a receptor regulatory protein	563	78
		norvegicus	precursor	1	ł
217	G04095	Homo sapiens	Hurnan secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus	Kupffer cell receptor	567	40
440	1 2003//	musculus	I repries con receptor	1 20.	"
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF238463 AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	636	96
222	AF021935		, , , , ,	030	90
	47 12 (607	norvegicus	kinase	(02	100
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
	1.5000135	 	11)	700	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus			
226	AE000218	Escherichia	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
	i	coli			L
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
		musculus			
229	AF122924	Xenopus	Wnt inhibitory factor-1	316	40
		laevis			
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific	290	100
255	1,7,7,7,7	110me saprems	phospholipase-D.		1
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
					62
237	X81466	Mus	Embryo Brain Kinase	460	02
026	11/64055	musculus	at all and at a DDVD of the control	204	1-2
238	U64857	Caenorhabditi	similar to the BPTI/Kunitz family of inhibitors;	284	33
	i	s elegans	most similar to tissue factor pathway inhibitor		1
	1		precursor (TFPI)		<u> </u>
239	AJ250840	Mus	serine/threonine protein kinase	739	63
		musculus	=:-		
240	AJ223472	Mus	transcription elongation factor TFIIS.h	222	38
		musculus			
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein	353	52
			sequence SEQ ID NO:18.		1
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-I	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
	1	norvegicus			1
244	AF016191	Rattus	potassium channel	1043	98
	1 2 3 3 3 3	norvegicus	poussium enumer	1015	1
			4:	645	98
	A E007344	Llomo sociore			. 70
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger		
245 246	Y29868	Homo sapiens	Human secreted protein clone pp325_9.	497	98
245 246 247	Y29868 AF180475	Homo sapiens Homo sapiens	Human secreted protein clone pp325_9. Not4-Np	497 188	98 83
245 246	Y29868	Homo sapiens	Human secreted protein clone pp325_9.	497	98

NO: 250 / 250 / 251 / 252 / 253 1 254 255 / 256 257 258 259 260 261 262 262 262 262 262 262 262 260 261 262 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260 261 262 260	AB022694 W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	sexta Kaposi's sarcoma- associated herpesvirus Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Fattus	protein SCLP Orf73 MOK protein kinase Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	Waterman Score 134 209 469 251 173 1201 460 368 1857 430	Identity
250 / 251 / 252 253 1 254 255 256 257 258 259 260 261 262 26	AB022694 W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Kaposi's sarcoma- associated herpesvirus Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens	MOK protein kinase Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	209 469 251 173 1201 460 368 1857	34 83 100 67 82 98 100 80
250 / 251 / 252 253 1 254 255 256 257 258 259 260 261 262 26	AB022694 W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Kaposi's sarcoma- associated herpesvirus Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens	MOK protein kinase Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	209 469 251 173 1201 460 368 1857 430	83 100 67 82 98 100 80
251	AB022694 W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	sarcoma- associated herpesvirus Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens	MOK protein kinase Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	209 469 251 173 1201 460 368 1857 430	83 100 67 82 98 100 80
252 253 1 254 255 7 256 257 2 258 2 260 7 262 7	W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ252968 AJ250839 AJ249977 AF141386 AF022859	sarcoma- associated herpesvirus Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens Homo sapiens	Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	469 251 173 1201 460 368 1857 430	100 67 82 98 100 80
252 253 1 254 255 7 256 257 2 258 2 260 7 262 7	W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ252968 AJ250839 AJ249977 AF141386 AF022859	herpesvirus Homo sapiens Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens	Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	469 251 173 1201 460 368 1857 430	100 67 82 98 100 80
252 253 1 254 255 256 257 2 258 2 260 2 261 2 262 2 2 2 2 2 2 2	W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ252968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	469 251 173 1201 460 368 1857 430	100 67 82 98 100 80
252 1 253 1 254 1 255 7 2 256 257 2 2 2 2 2 2 2 2 2	W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ252968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	469 251 173 1201 460 368 1857 430	100 67 82 98 100 80
252 1 253 1 254 1 255 7 2 256 257 2 2 2 2 2 2 2 2 2	W55045 L46815 W68505 AF070066 G02491 Z12841 Y95436 AJ252968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Neural adhesion molecule (ethb0018f2 product). DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	469 251 173 1201 460 368 1857 430	100 67 82 98 100 80
253 1 254 2 255 7 256 6 257 2 258 2 259 7 260 7 261 7	W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Mus musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	DNA binding protein Rc Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	251 173 1201 460 368 1857 430	67 82 98 100 80
254	W68505 AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	musculus Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Human acid sensing ionic channel. Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	173 1201 460 368 1857 430	82 98 100 80
255 // 256 (0 257 // 258 (2 259 // 260 // 261 // 262 //	AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	1201 460 368 1857 430	98 100 80 99
255 / 256 (257 258 259 / 260 / 261 / 262 / 262	AF070066 G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Mus musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Citron-K kinase Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	1201 460 368 1857 430	98 100 80 99
256 (257 2 257 2 258 3 259 4 260 4 261 4 262 4	G02491 Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	musculus Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6572. Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	460 368 1857 430	100 80 99
257 258 3 258 3 259 4 260 4 261 4 262 4	Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	368 1857 430	80 99
257 258 3 258 3 259 4 260 4 261 4 262 4	Z12841 Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	Oryctolagus cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Phospholipase Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	368 1857 430	80 99
258 259 A 259 A 260 A 261 A 262 A	Y95436 AJ222968 AJ250839 AJ249977 AF141386 AF022859	cuniculus Homo sapiens Mus musculus Homo sapiens Homo sapiens	Human calcium channel SOC-3/CRAC-2. L-periaxin serine/threonine protein kinase	1857 430	99
259 A 260 A 261 A 262 A	AJ222968 AJ250839 AJ249977 AF141386 AF022859	Homo sapiens Mus musculus Homo sapiens Homo sapiens	L-periaxin serine/threonine protein kinase	430	
259 A 260 A 261 A 262 A	AJ222968 AJ250839 AJ249977 AF141386 AF022859	Mus musculus Homo sapiens Homo sapiens	L-periaxin serine/threonine protein kinase	430	
260 A 261 A 262 A	AJ250839 AJ249977 AF141386 AF022859	musculus Homo sapiens Homo sapiens	serine/threonine protein kinase		12
261 A	AJ249977 AF141386 AF022859	Homo sapiens Homo sapiens		-	
261 A	AJ249977 AF141386 AF022859	Homo sapiens			100
262	AF141386 AF022859			861	100
	AF022859	i Kattus	AMP-activated protein kinase gamma 3 subunit	758	98
263			SLIT-2	198	40
263 I A		norvegicus			1
		Homo sapiens	neuropilin-2(a0)	335	62
	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
			(GPCR).		<u> </u>
266	U27269	Mus	sodium glucose cotransporter	204	56
		musculus			1
	AF 124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268 A	AF127389	Rattus	putative taste receptor TR1	209	39
}		norvegicus			j
	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
ł		pyogenes			
271	AB009883	Nicotiana	KED	109	26
		tabacum		1	1
272 A	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
		musculus			
273 I	L34938	Rattus	ionotropic glutamate receptor	460	86
1		norvegicus		1	1
274 A	AL022724	Homo sapiens	dJ413H6.1.1 (harnster Androgen-dependent	188	74
		-	Expressed Protein LIKE PUTATIVE protein)	Į.	
			(isoform 1)	1	1
275 A	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	173	94
		•	APOLLON		
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
	L40380	Homo sapiens	thyroid receptor interactor	430	61
	AB046851	Homo sapiens	KIAA1631 protein	283	96
	AC008075	Arabidopsis	Contains PF 00069 Eukaryotic protein kinase	157	43
		thaliana	domain.	'''	"
280 N	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
	AK024397	Homo sapiens	unnamed protein product	439	91
	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
/		musculus	210 domain adiooripuona repressor FE1	1 303	1 "
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate	647	100
-177	. 2,7,330	Trong sapiens	reading frame protein.	1 4,	100
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein	300	90
200	1 / 2402	LIOURO SUPIERS		300	1 70
286	A E() [64 1	Viene series	sequence SEQ ID NO:26.	 122	100
	AF016411	Homo sapiens	KCNA3.1B	137	100
	W89253	Homo sapiens	Human ALP.	688	97
	AF112886	Bos taurus	differentiation enhancing factor 1	750	96
	AF113131	Homo sapiens	host cell factor homolog LCP	367	44
	U52111	Homo sapiens	plexin-related protein	698	100
291 A	AF026504	Rattus	SPA-1 like protein p1294	603	89

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	<u> </u>
	1 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	norvegicus			<u> </u>
292	AF102854	Rattus	membrane-associated guanylate kinase-	124	53
293	X99211	norvegicus	interacting protein 2 Maguin-2	142	38
293	X99211	Drosophila melanogaster	ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	185	94
274	134343	rionio sapiens	sequence SEQ ID NO:92.	103	34
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577 1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	182	97
			sequence SEQ ID NO:92.		1
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	605	69
	}	•	gene 16 SEQ ID NO:63.	1	
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80_
305	305 U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
			Method: conceptual translation supplied by		Ì
			author	<u> </u>	
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
200	10005514	musculus			100
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
309	S79463	norvegicus	and the state of t	162	89
310	AF178941	Mus sp. Homo sapiens	semaphorin homolog=M-Sema F ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
311	003413	discoideum	calcium officing protein	131	30
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
	10.0	Zionio capiono	124 SEQ ID NO:124.	1	
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789 .	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
			44% similarity to U42767 (PID:g1736918)		
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and	278	38
			GENEWISE)	l	
316	U70209	Mus	polycystic kidney disease 1 protein	165	38
		musculus			
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
210	AF104923	norvegicus		1.20	\
318		Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma vivax	activated protein kinase C receptor homolog	141	38
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate	232	97
		Atomo supicits	receptor	1	1
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
		1		I	1
		1	123.	1	1
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
324 325	AF010144 M19650	Homo sapiens Homo sapiens		209 214	70 97
325			neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37)	1	
			neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214 140	
325 326 327	W80396 X75756	Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu	140 540	97
325 326 327 328	W80396 X75756 G02292	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373.	140 540 721	97 70 78 99
325 326 327 328 329	M19650 W80396 X75756 G02292 AF168990	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein	214 140 540 721 877	97 70 78
325 326 327 328	W80396 X75756 G02292	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein anti-HIV gp120 antibody heavy chain variable	140 540 721	97 70 78 99
325 326 327 328 329 330	M19650 W80396 X75756 G02292 AF168990 S67984	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein anti-HIV gp120 antibody heavy chain variable region	140 540 721 877 581	97 70 78 99 99 80
325 326 327 328 329 330 331	M19650 W80396 X75756 G02292 AF168990 S67984 X13916	Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein anti-HIV gp120 antibody heavy chain variable region LDL-receptor related precursor (AA -19 to 4525)	214 140 540 721 877 581 2823	97 70 78 99 99 80
325 326 327 328 329 330	M19650 W80396 X75756 G02292 AF168990 S67984	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein anti-HIV gp120 antibody heavy chain variable region LDL-receptor related precursor (AA -19 to 4525) Human signal peptide containing protein HSPP-	140 540 721 877 581	97 70 78 99 99 80
325 326 327 328 329 330 331	M19650 W80396 X75756 G02292 AF168990 S67984 X13916	Homo sapiens	neuronal thread protein AD7c-NTP 2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37) A secreted protein encoded by clone bp646_10. protein kinase C mu Human secreted protein, SEQ ID NO: 6373. putative GTP-binding protein anti-HIV gp120 antibody heavy chain variable region LDL-receptor related precursor (AA -19 to 4525)	214 140 540 721 877 581 2823	97 70 78 99 99 80

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
			similarity to P49205 (PID:g1345860)		
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	1111	67
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
352	D90777	Escherichia coli	activated potassium channel 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus norvegicus	phospholipase C delta-4	649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus	GTP binding protein	1456	91
		musculus			<u> </u>
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium	protein tyrosine kinase	115	44
		discoideum	 		
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates syndactylus	dopamine receptor D4	105	35
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1446	96
422	AC008075	Arabidopsis	F24J5.4	112	35
422					

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.) Species	Description	Waterman	Identity
NO:	110.	1		Score	lucinity
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ	1961	99
423	133342	riomo sapiens	ID NO. 191.	1901	1 99
426	AB009288	172	<u> </u>	635	98
426		Homo sapiens	N-copine		J
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
			associated protein.		
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon	extensin-like protein	613	48
		esculentum			1
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus	transmembrane receptor UNCSH1	817	93
		norvegicus			
450	AF081249	Homo sapiens	JAW1-related protein MRVIIA long isoform	4568	99
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
		Ì	(CIRP-1).		
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
	İ		gene 62.	i	
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum	,		ļ
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
			clone HTDAD22.		1
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
		1	gene 17.	1	
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein	135	47
			sequence SEQ ID NO:16.		
463	X84960	Triticum	low molecular weight glutenin	109	33
. 30].	aestivum			
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus	alpha/beta hydrolase-1	502	59
403	71.105/04	musculus	aiphia octa fryuroiasc-1	302	1 33
466	U93569	Homo sapiens	p40	101	30
				1	99
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	ود ا
460	000000	 	gene 77.	140	
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus	neurotoxin homologue	118	47
		musculus			
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
472	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	72	57

SEQ	Accession	Species	Description	Smith-	1%
ID	No.		S Compton	Waterman	Identity
NO:				Score	1dentity
			gene 62.		
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
	<u> </u>		sequence.		
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
478	G01870	17	clone H1DAD22.	-	<u> </u>
478	AF102777	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267 3427	100
4/9	AF102///	Mus musculus	FYVE finger-containing phosphoinositide kinase	3427	92
480	G03052	Homo sapiens	Human secreted protein, SEO ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
1.03		120mo sapiens	SYTAX1.	1 221	''
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486·	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
L			3	L	
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
493	J03799 U15174	Homo sapiens Homo sapiens	laminin-binding protein	228	70
473	015174	riomo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	197	67
	102050	Atomo supions	clone HTDAD22.	1 ''') 0,
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
		familiaris			į
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
500	U48508	maniculatus		25.05	
501	G03371	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
502	AF119851	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7452. PRO1722	105	58 62
503	AF113685	Homo sapiens	PRO0974	156 116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124 3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135 9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
			NO:180.		
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
			polypeptide kinase))		
509	U43360	Peromyseus	reverse transcriptase	97	62
510	C02790	maniculatus	II		
510 511	G03789 W79092	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
512	AF010144	Homo sapiens	Human secreted protein dn740_3.	1058	100
513	AJ133439	Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP GRIP1 protein	205	64
514	AE003456	Drosophila	CG6393 gene product	2151 259	100 42
J17	1 12003430	melanogaster	CO0373 Rette broduct	437	44 .
515	Z17206	Xenopus	p46XIEg22	128	40
	1	laevis	h r m Pre	120	70
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
518	AF151083	Homo sapiens	HSPC249	444	98
519	S80864	Homo sapiens	cytochrome c-like polypeptide	318	50
520	X92485	Plasmodium	pval	170	61
	<u> </u>	vivax			

SEO	Accession	Species	Description	Smith-	1%
ID	No.	Operics	Description	Waterman	Identity
NO:	1	ł		Score	120
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857	Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	253	73
		1	НРМВО32.		
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528	U47924	Homo sapiens	C8	1112	86
529	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	gin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi	contains similarity to a BR-C/TTK domain	853	39
240	711010430	s elegans	Contains similarity to a Divort 1K domain	855	"
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45%	408	66
541	ACOUSOS	Tionio sapiciis	similarity to P22059 (PID:g129308)	1 400	1 00
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus	olfactory receptor F3	327	73
5.5	7 102330	musculus	onacion receptor 13	321	1,3
544	Y73431	Homo sapiens	Human secreted protein clone yb186_1 protein	386	100
311	1,5451	Tromo suprems	sequence SEQ ID NO:84.	300	100
545	AE004833	Pseudomonas	probable TonB-dependent receptor	279	42
	}	aeruginosa	products roug approaches	}	l
546	G03793	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein	1772	67
	1	and the second	B receptor protein.		
548	Y91493	Homo sapiens	Human secreted protein sequence encoded by	176	100
			gene 43 SEQ ID NO:166.		
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584 2 protein	1224	94
			sequence.		* '
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
554	AB025258	Mus	granuphilin-a	501	41
		musculus).]]
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein	183	32
			complex component TRAP150	J ***	ļ
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus	Ca2÷ dependent activator protein for secretion	1010	93
		musculus	· Process of the control of		
561	AF187325	Canis	melanoma antigen	287	55
	1	familiaris]	
562	AJ001981	Homo sapiens	OXAIL	2512	99
563	Z17238	Rattus	glutamate receptor subtype delta-1	338	66
		norvegicus	O		"
564	W30638	Homo sapiens	Partial human 7-transmembrane receptor	371	100
	1	l rome supreme	HAPO167 protein.	}	} ***
565	AC005620	Homo sapiens	R33590 1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid	1138	78
200		J baprons	· · · · ·	1	1
200	i		sequence SEO ID NO:63.		1
567	AL031177	Homo sapiens	sequence SEQ ID NO:63. dJ889M15.3 (novel protein)	1002	58

SEO	Accession	Species	Description	Smith-	1%
ID ID	No.	Opocies	Description	Waterman	Identity
NO:	110.	1		Score	lucitoria
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
	Y07096			1064	100
571	107096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
550	41.021155	1	sequence.	L	1
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
			ID NO:388.	1	
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
		familiaris			1
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
303	7.2030020	110mo suprans	Antigen)	1 200	03
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
203	AF253017	musculus	2F1 protein	704	80
590	11/00/07		Secreted protein encoded by gene 94 clone	220	81
390	W88627	Homo sapiens		329	01
	V20000	<u> </u>	HPMBQ32.		-
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
500	1152055	 	protein.	1000	
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
500	-	 	polypeptide.		ļ
593	Y53051	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	97
		 	sequence SEQ ID NO:108.		<u> </u>
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus	COP1 protein	2215	95
		musculus			
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus	putative secreted protein ZSIG37	143	40
	<u> </u>	musculus			<u> </u>
599	AF119855	Homo sapiens	PRO1847	236	76
600·	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
_	ļ		Antigen)		ļ
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756 ·	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
		l secure capions	HPMBQ32.]
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
	1	ouploid	107.	[
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
013	102033	Tionno sapiens	clone HTDAD22.	190	"
614	M87053	Rattus		450	84
014	MO 1033	norvegicus	lens membrane protein	450	04
615	A CO0 4222		ED14216	162	37
615	AC004232	Homo sapiens	FPM315	163	37
616	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapicns	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	T **	Homo sapiens clone BK158_1 protein.	1202	99
		Homo sapiens		1203	77
	AF257330	Homo sapiens	COBW-like protein	1440	98
649	AF257330 Y36203	Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75.	1440 233	98 73
650	AF257330 Y36203 G02872	Homo sapiens Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953.	1440 233 173	98 73 78
649 650 651	AF257330 Y36203 G02872 Y32199	Homo sapiens Homo sapiens Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379.	1440 233 173 1012	98 73 78 100
649 650 651	AF257330 Y36203 G02872 Y32199 AB032909	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4	1440 233 173 1012	98 73 78
649 650 651 652	AF257330 Y36203 G02872 Y32199 AB032909 AK021848	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product	1440 233 173 1012 122	98 73 78 100
649 650 651 652 653 654	AF257330 Y36203 G02872 Y32199 AB032909	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No. 15.	1440 233 173 1012	98 73 78 100
649 650 651 652 653 654	AF257330 Y36203 G02872 Y32199 AB032909 AK021848	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No.	1440 233 173 1012 122	98 73 78 100 32
649 650 651 652 653 654	AF257330 Y36203 G02872 Y32199 AB032909 AK021848 W73411	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens Homo sapiens Rattus	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No. 15.	1440 233 173 1012 122 186 57	98 73 78 100 32 69 37
649 650 651 652 653 654 655 656 657	AF257330 Y36203 G02872 Y32199 AB032909 AK021848 W73411 L22455	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens Homo sapiens Rattus norvegicus	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No. 15. mu opioid receptor	1440 233 173 1012 122 186 57	98 73 78 100 32 69 37
649 650 651 652 653 654 655 656 657 658	AF257330 Y36203 G02872 Y32199 AB032909 AK021848 W73411 L22455 G03112 G02345 W88627	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No. 15. mu opioid receptor Human secreted protein, SEQ ID NO: 7193. Human secreted protein, SEQ ID NO: 6426. Secreted protein encoded by gene 94 clone HPMBQ32.	1440 233 173 1012 122 186 57 116	98 73 78 100 32 69 37 34
649 650 651 652 653 654 655 656 657	AF257330 Y36203 G02872 Y32199 AB032909 AK021848 W73411 L22455 G03112 G02345	Homo sapiens Homo sapiens Homo sapiens Homo sapiens Hylobates agilis Homo sapiens Homo sapiens Rattus norvegicus Homo sapiens Homo sapiens	COBW-like protein Human secreted protein #75. Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by Incyte clone 2022379. dopamine receptor D4 unnamed protein product Human secreted protein encoded by Gene No. 15. mu opioid receptor Human secreted protein, SEQ ID NO: 7193. Human secreted protein, SEQ ID NO: 6426. Secreted protein encoded by gene 94 clone	1440 233 173 1012 122 186 57 116	98 73 78 100 32 69 37 34 45

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	1	l		Score	lacinity
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	375	43
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein (hs6M1-21))	480	55
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KJÅA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88 .
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1 .		Waterman	Identity
NO:		<u> </u>		Score	ļ
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 2DD).	121	95
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	1 59
,,,	1.205577	es cerevisiae	67.7	12.	1 "
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
		norvegicus			
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota	olfactory receptor	615	83
	 	marmota			·
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380 80	100
716 717	G00577 Y96864	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4658. SEQ. ID. 37 from WO0034474.	835	73
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
	077337	170mo supiens	receptor beta4 subunit	3,0	"
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
			designated HSCOP-6.		
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter,	111	41
		tigrinum	NBC		
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
725	X54673	musculus Homo sapiens	protein 3A GABA transporter	3114	99
726	AF016191	Rattus	potassium channel	370	100
720	A1 010191	norvegicus	potassium channer	370	100
727	ΔΒ029559	Rattus	BATI	139	35
		norvegicus		}	1
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila)	142	68
	7:000	 	homolog)		ļ
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor	675	99
732	AF161382	Homo sapiens	homologue Vanilrep1. HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia Escherichia	putative transport protein	592	97
	1.2000.33	coli	parative damoport protoni	""	1
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	2173	99
		•	(rhodopsin family) protein similar to high-]	}
			affinity lysophosphatidic acid receptor homolog)		<u> </u>
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
222	V55010	 		003	1.00
737 738	X55019 X91906	Homo sapiens Homo sapiens	acetylcholine receptor delta subunit	883 1978	99.
738	AB026116	Homo sapiens	voltage-gated chloride ion channel organic anion transporter 4	1978	100 98
740	D00570	Mus	open reading frame (196 AA)	83	24
,40	1 2003/0	musculus	open reading traile (170 AA)	١ "	27
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32	448	95
			clone HLTCJ63.		
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA	968	65
740	V04025	1,12	from plasmid pGCS2232.	632	100
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein	622	100
749	AL022238	Homo sapiens	sequence SEQ ID NO:76. dJ1042K10.5 (novel protein)	314	85
744			COLOTEINION HIDYGI DIVIEIII)	, ,14	1 03

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus musculus	granuphilin-a	773	41
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus	vasopressin receptor	979	68
	14105105	norvegicus			
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8B192.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 della	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOVI	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
795	Y69384	Homo sapiens	Amino acid sequence of a 14274 receptor protein.	119	100
	W40215	Homo sapiens	Human macrophage antigen.	1358	99

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.	Species	Description	Waterman	Identity
NO:	140.			Score	lacinary
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
				1913	92
801	L00073	Homo sapiens	renin		1
802	P92219	Homo sapiens	CR1 protein.	11963	97
	<u> </u>	(human)			<u> </u>
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
	1		LAT2		1
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
			clone HOVBA03.		ł
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
0.5	711 203772	Treme suprems	encoded by GenBank Accession Number	,,,,	***
i	1	1	L25899		
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
810	W 93030	Homo Sapiens	gn114 1.	336	100
817	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
			CGI-41 protein		95
818	AF151800	Homo sapiens		1106	
819	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
820	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored	4897	99
		<u></u>	protein GPI-122.		
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma-	1105	100
	<u> </u>		2 subunit		
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded	1540	100
			from gene 28.		
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
			gene 24 SEQ ID NO:147.		
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262 .	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828_1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi	glycine-rich	85	36
		s elegans		!	1
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in	998	75
			AL023803))		
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543_3.	1105	67
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
842	AF036717	Homo sapiens		2629	99
	1		FGFR signalling adaptor SNT-1		
843	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein	1089	100
044	COSCO	Via-re-	sequence SEQ ID NO:114.	257	
844	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	357	69
845	AF151810	Homo sapiens	CGI-52 protein	1443	88
846	X83378 AC004883	Homo sapiens Homo sapiens	putative chloride channel similar to general transcription factor 21; similar	1620	99
847			consider to company termination factor 11, rimilar	655	96

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:		 	to AE028060 (DID) -2822207)	Score	
848	X99886	Homo sapiens	to AF038969 (PID:g2827207) monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
850	AB038237		P34984 (PID:g464305)	1767	
851	AF124490	Homo sapiens	G protein-coupled receptor C51.2	3415	100
	d	Homo sapiens	ARF GTPase-activating protein GIT1		98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1048	98
871	L04311	Homo sapiens	gene 25 SEQ ID NO:305. GABA-alpha receptor beta-3 subunit	237	
872 ·	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	93
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412			464	100
875	Y27572	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7493. Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	'Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted	321	100
880	W88991	VIama saniona	protein.	026	100
881	AF118670	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
882	AF208865	Homo sapiens Homo sapiens	orphan G protein-coupled receptor EDRF	1971 528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein	348	100
885	AF070661	Homo sapiens	sequence SEQ ID NO:106. HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887	X92744	Homo sapiens	hBD-1	375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone	994	94
889	Y41293	Homo sapiens	cn621_8. Human soluble protein ZTMPO-1.	4595 ·	99
890	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo sapiens	G0S19-2 peptide precursor	481	100
897	Z68747	Homo sapiens	imogen 38	2018	99
•			neurokinin B-like protein ZNEUROK1	619	100
898	AF186112	Homo sapiens	neurokinin B-like protein ZNELIRLIK I		

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO: 900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
900	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
903	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688		antigen NY-CO-3	771	99
		Homo sapiens		2544	100
906	AB007836	Homo sapiens	Hic-5		
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162 1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
				T	98
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
		familiaris			
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
943	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ	667	100
			ID NO. 463.		
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ	Accession	Species -	Description	Smith-	1%
ID `	No.	1		Waterman	Identity
NO:				Score	
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
			ID NO. 496.		
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726			1879	100
		Homo sapiens	Human secreted protein fg949_3.		
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein sequence SEQ ID NO:110.	587	100
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CG1-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene:Acc#	1466	100
			AF030433		
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	mcmbrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
		•	gated cation channel hHCN4		
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261 .	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP80	3348	99
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
980	AF119297	Homo sapiens	novemendanina manifia metain like metain 1	1170	99
981			neuroendocrine-specific protein-like protein 1	1983	99
	AF155652	Homo sapiens	potassium channel modulatory factor		1
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412- cncoded protein.	1553	99
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	172	70
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca sexta	juvenile hormone esterase binding protein	226	32
989	G03697		Human secreted protein SEO ID MO: 7770	194	88
999	AF204159	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	1486	100
		Homo sapiens	potassium large conductance calcium-activated channel beta 3a subunit		
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi s elegans	VM106R.1	327	40
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
95	AF133845	Homo sapiens	corin	5811	99
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein	4999	100
007	WCOCC	IIama	complex component TRAP150	204	102
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID NO:212.	725	100
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56.	676	47
1001	AF190167	Homo sapiens	membrane associated protein SLP-2	1747	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005 ·	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010 ·	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016 ·	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMP1	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040 -	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049 ·	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135 Y41674	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	AJ250042	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	Y36087	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575 770	100 85
·		Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.		
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072 1073	AL031177 X82200	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1074	G03213	Homo sapiens Homo sapiens	gpStaf50 Human secreted protein, SEQ ID NO: 7294.	249 99	62 47
1075	Y36233	Homo sapiens	Human secreted protein, SEQ ID NO. 7294. Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080 ·	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090 1091	G04063 S72304	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
		Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	149	59

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	-		Waterman	Identity
NO:				Score	
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus musculus	ribosomal protein L28	128	69
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 155.	244	97
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by gene 49 SEQ ID NO:170.	542	100
1134	AB017908 ·	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91 ·
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis familiaris	D4 dopamine receptor	89	48
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:		s elegans	cerevisiae zinc resistance protein	Score	ļ
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 7519. Human secreted protein, SEQ ID NO: 5084.	181.	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 384. Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
1154	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens		341	87
1156	Y86265	Homo sapiens	Tumour suppressor protein, p53. Human secreted protein HUSXE77, SEQ ID	99	41
			NO:180.		
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocystcine associated protein	249	62
		norvegicus	•		
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	107	71 .
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens C1542 2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Hurnan secreted protein HELHN47, SEQ ID NO:175.	1089	89
	l-	1	110.173.	I	L

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	•	•	Waterman	Identity
NO:				Score	
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203 ·	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	polypeptide. High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB). Extended human secreted protein sequence, SEQ	135	31
1000	1,000	<u> </u>	ID NO. 160.		ļ
1223	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226 1227	G01733 AF099973	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	333	100 56
		Mus musculus	schlafen2		
1228	G01218 AF217188	Homo sapiens Mus	Human secreted protein, SEQ ID NO: 5299.	155 801	63
		musculus			
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231 1232	X98333	Homo sapiens	organic cation transporter	1704 212	100
1232	W74955 Y94940	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24. Human secreted protein clone yi62 1 protein	526	100
	1		sequence SEQ ID NO:86.		
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236 1237	G01459 AF000018	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 5540. adapter protein	417 164	100
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34 1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1,240		, unpress	Extended human secreted protein sequence, SEQ		

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	<u> </u>
			ID NO. 160.		
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
		norvegicus	<u> </u>	<u> </u>	
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	<u> </u>	<u> </u>	gene 27 SEQ ID NO:131.	<u> </u>	
1251	L15313	Caenorhabditi	putative	858	59
	<u> </u>	s elegans		<u> </u>	<u> </u>
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate	278	75
			reading frame protein.	<u> </u>	
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-1	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
			from gene 26.		
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	986	100
			ligand (clone 3TW).		
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	288	71
			sequence.		
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
		İ	ID NO:2.		
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus	phosphatidylinositol 5-phosphate 4-kinase	859	95
		norvegicus	gamma		
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2	552	76
		musculus		ĺ	
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
			GTPase-activating protein p135 SynGAP)]	
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform .	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme	1280	100
	1		similar to acetyl-coenzyme A synthethase	ļ	
	l		(acetate-coA ligase))	1	1
1274	AF064748	Mus	S3-12	3523	61
		musculus			
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted	643	90
			protein.		1
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus	aorta CNG channel (rACNG)	267	85
		cuniculus			
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1283	AF117814	Mus	odd-skipped related 1 protein	357	98
		musculus			
1284	U87318	Xenopus	NaDC-2	535	60
		laevis			
1285	AF061346	Mus	Edp1 protein	452	68
		musculus			1
1286	AB030182	Mus	contains transmembrane (TM) region	582	68
		musculus	\		1
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
		construct			-
1288	AF254411	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98
1289	(Aruo42U)	: Italius			

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1	-	Waterman	Identity
NO:				Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339 ·	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1	1351	A	2	337	1	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC HWPQAPHRA***GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMRQMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R SCNYTLALILFL
4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł	•	peptide	ł	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
						WSHEGEILQAFRGHQGRGIRAIAAHERQAWV
	1	ļ			j	ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ
i	ļ	İ				VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
			1		1	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
		1				RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
						WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
				ļ		LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
)		Ì	ļ	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
						QCLGFVDSDSRKMVSTLT
9	1359	Α	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN
	1	<u> </u>	-	<u> </u>	1046	KSSEFNEGPERERMDV
10	1360	Α	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY
	[1	Į.		1	FEEVQRLRFEVHDISSNHNGLKEADFLGGME
Ţ	ĺ	İ				CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA
	1	1	Į	ļ	ļ	EELSGNDDYVELAFNARKLDDKDFFSKSDPF
	ł	l	i	}	}	LEIFRMNDDATQQLVHRTEVVMNNLSPAWK
						SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK
				1		HDFIGEFTSTFKEMRGAMEGKQVQWECINPK
İ	[}				YKAKKNYKNSGTVILNLCKIHKMHSFLDYI
1			1			MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP
	ļ					YQPNEYLKALVAVGEICQDYDSDKMFPAFGF
1	i	i				GARIPPEYTDSHDFAINFNEDNPECAGIQGVV
			l			EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR
	1	1)			SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP
	1221-		145		9	DNPGGHFV
11	1361	A	147	614	٩	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT
	•					NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL
						KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS
	ĺ	ĺ	ĺ			
		İ	[DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS
			Ì			TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI
12	1302	^	'''	12	710	DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM
1	ļ	}	l i			VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA
		ŀ				ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI
1	1	ļ				NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ
1	****	١				MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS
1	ĺ	ĺ	[TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI
1	[(DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC
1			[QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH
l	1304	l '`		3.2		FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT
ĺ])			*SFFFIFSWGTNGCLLSAITYACYAAICHPLLS
		1				TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM
	1	1			1	AIEFLLECDONIT\KLICENT*KNIAKNI*KRRV
[[1				TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI
	1	1				KOTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK
	}	, ,	203		.01	SKDVFSKPVNIFWALEESVLGVKARQPKPFFA
	1					AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV
	ļ	l				GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW
l	1.50,	١.,		-		VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL
1.0	1	l	"""		· ·	ELLTSGDPPALASQSAGITGMSHCARPKGHFG
	Ŀ	Ь		L		PETTODI I UTUDANU I OMNITOMI KOULA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHLLSHIWKDVQNILLK
20	1370	A	304		1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW
21	1371	A	326	799	1587	ALVRSRGG GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	Α	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ
25	1375	A	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	Α	397	383	165	EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/I/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRP\ANF*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	A	474	125	471	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825		KSACSFICSEEQPASPSPLKPGTYASETRPRDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGGVETEGAG*CPAAPGTDLR APGGSAGS*VGLPSAGGSRGKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRQE*P PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS* GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	. 512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRPVVPFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

NO. of No. of N	SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuclecutic seq						1	
Sequence	1		1.00				
Sequence			l				
1914 ng to first simio acid residue of peptide pep			l				
minino acid residue of peptide residue of peptide residue of peptide sequence T-Tinteonine, V-Valine, W-Tyrophan, Y-Sup codon, Peposible nucleotide deletion, I-possible nucleotide deleti		dence	i	1			
Peptide Sequence	uence		j	914			
Peptide							
						sequence	
GRPSQSOPAGPRIMRGCCLRGW*PSSSGSD GGPHPASTWLRAGKTGPSPACOCA*IPPPS VSAAPQSPETECPRGCAAAAGCVLAAAGGS HAGIGLPGVRVHTGRGGCTPPR LRSI_PVLGI_PAPRCPVSAHPWHRSGSCCIAA ARLYPREPAPCGP**IPPA*GCP** LRSI_PVLGI_PAPRCPVSAHPWHRSGSCCIAA ARLYPREPAPCGP** TOPPITTION TOPPITTI	}	1	l	ļ	peptide		/=possible nucleotide deletion, \=possible
GPGPHPASTWLRACKTGPSPPACCCA*LPPPS		1		}	sequence		nucleotide insertion
GPGPHPASTWLRACKTGPSPPACCCA*LPPPS			 				GRPSOSO*PAGPPGMRGCCLRGW*PSSSGSD
			ľ	1			
HGAGIJPOVRVHTORVHHIPGAGGCCTPRED		1		1	[
LRSLPVLGLPARRCYSAHPWHRRSGSSCH ARLVPRIEARGCY-TGVPLTGPPT-9 GLP NHOAVGLEASGALQAGHRDELPTMVQLLDH SPYPPLKGRPHAP SPYPPLKGRPHAP FRLPLAAGA/RGAAEPRAVSMAPDPSAKIH WEASPEMGSKCHOKGRNOŢIECFNITVRITJ RLSPLAAGA/RGAAEPRAVSMAPDPSAKIH WEASPEMGSKCHOKGRNOŢIECFNITVRITJ FRLPLAAGA/RGAAEPRAVSMAPDPSAKIH WEASPEMGSKCHOKGRNOŢIECFNITVRITJ FRLPLAAGA/RGAAEPRAVSMAPDPSAKIH WEASPEMGSKCHOKGRNOŢIECFNITVRITJ FREGKEKCPYDPARGFIGLIDGGLYTATVS FREGKEKCPYDPARGFIGLIDGGLYTATVS FREGKEKCPYDPARGFIGLIDGGLYTATVS FREGKEKCPYDPARGFICLIDGGLYTATVS FREGKEKCPYDPARGFICLIDGGLYTATVS SIPVCQVRGQPGSGKESPACLKSLSCICTH NDAEFVFSVLVRESKASAVGDDDKYYYFFTE RATEKEGSSTIGSSIRVARGPIC RATEKEGSSTIGSSIRVARGPIC SIPVCQVRGGCGGARAEVQOGONRGKG TSEGKEG-PYDVPLPVPSPLPPLPCJCMLDYL KDKKEVGFFGSIQALMQTICGEKVMADDEFT VKDKEVGFFGSIQALMQTICGEKVMADDEFT NDAEFVLOLLEGEINNDFQNILKTQTGPTT NIHICTVDYLLRLGESI TUDITGPILLIGGVPJVPKDFRGRNRQFGGCM RNLSVDGKNVDMAGFIANNGTRECCAARN FCDGRRRQNGGTCVNRWMMLCECCLRROG KNCEGGEWPASSIPVTAAWEALLDLVPGTT VRGILHIQVRQPLVYVAAFTVDSIRFLIGETVL RRAPASGOYPSSEVOWDR AGPAEPSPSTP ATVIISVPWYLGIMFRTRKEDSVLMEATSGE FTSRLQVTGARCHGOTCVGARGPAEPSPSTP ATVIISVPWYLGIMFRTRKEDSVLMEATSGE FTSRLQVTGARCHGOTCVGARGH FRLDTAGG	1	i		1			
ARLVPRIPAGCC**TG*VILTGFPEP*AGLDNNQLDH NHQAVGLASGALQAGIRDELPTMVQLLDH SPDYPLKGRPHAP 37		j		}	j	į.	,
				1			
SPDYPLKGRPHAP					1		
1387 A 620 828	l	i	l	i	ì	ì	
WEASPEMOSKCHOKGKNNOTECFNHYRFLQ RIANSTHLYACGTHAPPUL-ADAFATLETS FEEGKEKCPYDPARGFTGLIDGGLYTATRYE FRSIPDIRRSRHPHSLRTEETFMHWLNG*EDE AQDDGG*GTISSFLLPWADIDAFATLETS FEEGKEKCPYDPARGFTGLIDGGLYTATRYE FRSIPDIRRSRHPHSLRTEETFMHWLNG*EDE AQDDGG*GTSSFLLPWADIDAFEDE ADDGGG*GTSSFLLPWADIDAFEDE SPVCCQVRGQPQSGGKESPACLKSLNCLTH UDAFFYSFLVLYRSKASAVGQDEVYYFFTE RATEKESGSFTQSRSSHRVARGDPPL RATEKESGSFTQSRSSHRVARGDPPL GENEVALOGEVGVAQDEKGG TSEEGKEG*EVPV*LPVSPPLPPLQKMLDYL KDKKKVGFFQSIQALMQTCGEKVVAADDEFT QDLFRFLQLLCEGHNNDFGVJKRTQTGNTTT NIIICTVDYLLRLQESI TOLTOPLLLGGVNNVFDFRGRNRQFGGCM RNLSVDGKNVDMAGFIANNGTREGCAARN FCDGRRQNGGTCVNRWNYLCECPLRFEGG KNCEQGEWPASSIPPVTAAWEALLLDVPGTT VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL RRAPAPASGVSPSFGVGWSFAGEPSSTST ATVIISVPWYLGLMFRTRKEDSVLMEATSGG PTSFFRLQVTGAPCHQGTCVGRQFDMLSG LRVTDGEWHHLLIELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWGTPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL TLDYGMDQVSWHLHLLWGTPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL TLDYGMDQVSWHLHLLWGTPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL TLDYGMDQVSWHLHTLWGTPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL TLDYGMDQVSWHLHLLWGTVYYQA VNRSVVLQYTIM TLDYGMDQVSWHLHTLWGTPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL TLDYGMPGVSVMSLAKRRVY QGGTVASNLSQL*TLNAFFPELLFRSLARTG FVLT-NRFGPGLYTWYGFGLGCHINWTGFGLGCHINWTGFGLGCHINWTGFGLGCHINWTGFGLGCHINWTGFGLGCHINWTGFGLGCHINGT QGGTVASNLSQL*TLNAFFPELLFRSLARTG FVLT-NRFGPGLYTYWGFGLGCHINGT CAPPRL*PRRGGDPGLDCNRRGILL KACFFTINVTL QRFMLPPSHAGLARPPPPEISVF PLT-NRFGPGLYTWYGFGLGCHINFWES PRT-MLRSASQDLNAFTLVGHPSRFLQG QVSCPPQFTLPREKFLLHRPPRPAQPPLPR PLT-STRRNUDPEIPER PLT-ST		1		1			SPDYPLKGRPHAP
RINSTHILVACGTHAFQPLCAAIDABAFTIPTS FEEGKEKCPYDPAAGFTGIDCTYATRYE FRSIPDIRRSRHPHSLRTEETPMHWLNGFEDE AQDDGG*GTISSTLPWADHFTPKSPCIPVH SIPVCQVRGQPQSGKESFALKSLSNCLTH DAEFVPSVLVRESKASAVGDDLKYYYFFTE RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL REMAVSSTLKLGISILNGGNAEVQQORINGKG TSEGKEG*EYPV*LPVSPIPPLQKMLDVL KDKKEVOFFQSIQALMQTCGEKVMADDEFT QDLFRFLGLLGEGHINDPQNYLRTQTGITNTT NIIICTVDYLLRLQESI NIIICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI RINISYDGKNVDDMAGFIANDGTBGCAARRN FCDGRRQNGGTCVRRWNNYLCECPLRFGG KNCEQGEWRASIPPVTAWGHIQVQPLVVYAAFTVDSHRPLQETVL RRAPAPASGVPSPSGVGWBY APAPERSSTP ATVIISVPWVLGIMFRTRKEDSVLMEATSGG PTSFFLQVTGAFCHQGTGVGARDPMLSG LRVTDGEWHHLLELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWG*TLPFAQGKTGA SEDKVSVRRGFGGCMGACPS SEDKVSVRRGFGGCMGACPS GVAPRL RQLV-PLSTC*GVKYLLDAVHTWTYEAR VNSVVLVQVTIM TLDYGMDQVSWHLHLIWG*TLPFAQGKTGA SEDKVSVRRGFGGCMGACPSGGGGGGGACPS SEDKVSVRRGFGGCMGACPS GVAPRL RQLV-PLSTC*VKYLLDAVHKTYYEAR VNSVVLVQVTIM SIMLKERKVFQFPSCLFFQYTTWLGFPYHVLFD SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT GGTHASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QGGTYASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QGGTYASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QRFML*PSHAGLARTP*PEPILVF PLT*NRFRFQLTYFRSKNLVHLARL**PSTALK*PTR SPRCIPRTSLYFTICO*PPALQL*SPIEPPATYR PPT*MLRSASQPLNQAFTLVGHPSPRTQG QVSCPPQFTLPRSKRLVHLRAFSLAFTPTATYR PPT*MLRSASQPLNQAFTLVGHPSPRTQG QVSCPPQFTLPRSKRLVHLALT**PSTALK*PTR SPRCIPRTSLYFTICO*PPALQL*SPIEPPATYR PLT*STRRND*PDEPIERFR PT*MLRSASQPLNQAFTLVGHPSSFQTYGT ULSMSMGK EQXYMTROPC*VTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CV	37	1387	Α	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
RINSTHILVACGTHAFQPLCAAIDABAFTIPTS FEEGKEKCPYDPAAGFTGIDCTYATRYE FRSIPDIRRSRHPHSLRTEETPMHWLNGFEDE AQDDGG*GTISSTLPWADHFTPKSPCIPVH SIPVCQVRGQPQSGKESFALKSLSNCLTH DAEFVPSVLVRESKASAVGDDLKYYYFFTE RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL REMAVSSTLKLGISILNGGNAEVQQORINGKG TSEGKEG*EYPV*LPVSPIPPLQKMLDVL KDKKEVOFFQSIQALMQTCGEKVMADDEFT QDLFRFLGLLGEGHINDPQNYLRTQTGITNTT NIIICTVDYLLRLQESI NIIICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI TRIMICTVDYLLRLQESI RINISYDGKNVDDMAGFIANDGTBGCAARRN FCDGRRQNGGTCVRRWNNYLCECPLRFGG KNCEQGEWRASIPPVTAWGHIQVQPLVVYAAFTVDSHRPLQETVL RRAPAPASGVPSPSGVGWBY APAPERSSTP ATVIISVPWVLGIMFRTRKEDSVLMEATSGG PTSFFLQVTGAFCHQGTGVGARDPMLSG LRVTDGEWHHLLELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWG*TLPFAQGKTGA SEDKVSVRRGFGGCMGACPS SEDKVSVRRGFGGCMGACPS GVAPRL RQLV-PLSTC*GVKYLLDAVHTWTYEAR VNSVVLVQVTIM TLDYGMDQVSWHLHLIWG*TLPFAQGKTGA SEDKVSVRRGFGGCMGACPSGGGGGGGACPS SEDKVSVRRGFGGCMGACPS GVAPRL RQLV-PLSTC*VKYLLDAVHKTYYEAR VNSVVLVQVTIM SIMLKERKVFQFPSCLFFQYTTWLGFPYHVLFD SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT GGTHASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QGGTYASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QGGTYASNLSQL*TLNAFPFELLLFRSLARTG FVL*NRFGFGLVTYWFGFQELDCRREGGLL KACFFTNIVTL QRFML*PSHAGLARTP*PEPILVF PLT*NRFRFQLTYFRSKNLVHLARL**PSTALK*PTR SPRCIPRTSLYFTICO*PPALQL*SPIEPPATYR PPT*MLRSASQPLNQAFTLVGHPSPRTQG QVSCPPQFTLPRSKRLVHLRAFSLAFTPTATYR PPT*MLRSASQPLNQAFTLVGHPSPRTQG QVSCPPQFTLPRSKRLVHLALT**PSTALK*PTR SPRCIPRTSLYFTICO*PPALQL*SPIEPPATYR PLT*STRRND*PDEPIERFR PT*MLRSASQPLNQAFTLVGHPSSFQTYGT ULSMSMGK EQXYMTROPC*VTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CVTHDISLYIVTKLALIFLIPPT LEQXYMRTDO*CV				ſ	[1	WEASPEMOSKCHOKGKNNOTECFNHVRFLO
FEEGKEKCPYDPARGTGILIDGGLYTATRYE		ſ		1		1	1
FRSIPIRRSRIPHSLRTEETPMIWLNO*EDE			1	1		ł	,
AQDDGG*GTISSFILPWPADIPTPKSPGEPVH SIPVCQVRGQPQSGKESPACLKSLSNCLTH				1	1	1	
SIPVCCQVRGQPOSGRESPACLKSLSNCLTH		ŀ	l		•		
				1	ļ		
RATEKESCSFTQSRSSHRVARGIPPL		İ		Î	Í	l	
1388			١.	1			
TSEEGKEG*EVPV*I_PVSPPLPRPLOKMLDVI. KDKKEVG*FFQSIQALMQTC\GEKVMADDEFT QDL*FRFLQLLCEGHNNDFQNYLRTQTGNTTT RNIIICTVDYLLRLQESI		,		}			RATEKESGSFTQSRSSHRVARGIPPL
SUNKEVGFFGSIQALMQTCGEKVMADDEFT QDLFRFLQLLCEGHNNDFQNYLRTQTGNITT NIIICTVDYLLRLQES SULFICIAL SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC Sulfic S	38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG
SUNKEVGFFGSIQALMQTCGEKVMADDEFT QDLFRFLQLLCEGHNNDFQNYLRTQTGNITT NIIICTVDYLLRLQES SULFICIAL SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC SULFIC Sulfic S		1	İ	1			TSEEGKEG*EVPV*LPVSPPLPRPLOKMLDYL
ODLFRFLQLLGEGHNNDFQNYLRTQTGNTTT NINICTVDYLLRLQES 39		l	1	ł	1	1	
1389		•	ŀ	ł	ł	1	
1389		l .		ļ			
RNLSVDGKNVDMAGFIANNGTREGCAARRN FCDGRRRQNGGTCVNRWNMYLGECPLRFGG KNCEQGEWPASSIPPVTAAWEALLLDVPGTT VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL RRAPAPASQVPSPSGVGWDR*AGPAEPSPSP ATVIISVPWYLGLMFRITRKEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VQARGRDPMLSG LRVTDGEWHHLLELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL 40 1390 A 801 69 399 HIKIIHKEDLNKWKYILCSGMERLSTVMIPVV PQIIYKFNA*QVILKFTW*E*GAKITILRKNKL RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR VNKSVVLVQVTIM VNKSVVLVQVTIM SSMLKERKVFQFFSCLFFQYITWLGPPYHVLFD SSVNPSIGAK*DILQSVMNCLYAKRIPCVT GSTHASGYDKTPDFILQVPVAVEGHIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLATIG FVLT*RFGPGLVIYWYGFIQELDCNRERGILL KACFPTINIVIL KACFPTINIVIL KACFPTINIVIL QRPMLPPSHAGLARPPPEPISVP A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPFHTLLP QRPMLPPSHAGLARPPPEPISVP GVGPTNSKSLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQLSPEEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPEFRR QEFYDRLRTLS*PQTSIGNLEFPIYGT WLSMSMGK TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMKT/DQ*CVTHDISLYIVTKLALIFLIPR EIQKYMK	20	1200		7/7	 	1020	
FCDGRRQNGGTCVNRWNMYLCECPLRFGG KNCEQGEWPASSIPPVTAAWEALLDVPGTT VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP ATVIISVPWYLGLIMFRTRIKEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG LRVTDGEWHHLLIEKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLWG*TLPPAQGKTGA SEEDK VSVRGFRGCMQVRGGCGGRGEACPS QAAPRL	39	1389	Α	/6/	1	1030	
KNCEQGEWPASSIPPYTAAWEALLIDVPGTT		<u>l</u>		ł	ļ		
VRGLHIQVRQPLVYYAAFTVDSHRPLQETVL RRAPAPAGVPPSPSGVGWDR*AGPAEPSPSTP ATVIISVPWYLGGMFRTRKEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG LRVTDGEWHHLLILEKNVKEDSEMKHLVTM TLDYGMDQVSWHLHILLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL 40 1390		1	ł	ł	ł		1
RRAPAPASGVESPSGVGWDR*AGPAEPSSTT ATVIISVPWYLGLMFRITRLEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VQARGRDPMLSG PTSFRLQVTGAPCHQGTC*VQARGRDPMLSG LRVTDGEWHHLLIELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL		1		ļ			KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
ATVIISVPWYLGLMFRTRKEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VQARGRDPMLSG LRVTDGEWHHLLELKNVKEDSEMKHLVTM TlDYGMDQVSWHLHLLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL 40		1		-	-		VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG	1			İ	ĺ		RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG		l	İ	Į			ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
LRVTDGEWHHLLIELKNVKEDSEMKHLVTM		ļ	j	J	}	}	1
TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGGEACPS QAAPRL	1	ŀ]	J :	1	}	
SEDKVSVRRGFRGCMQVRGGCGGRĞEACPS QAAPRL	1	ļ		1		ļ	
QAAPRL	1	1					
1390		l		i I		İ	
PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR VNKSVVLVQVTIM			<u> </u>				
RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR	J 40	1390	A	801	69	399	
VNKSVVLVQVTIM			l				
41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT 42 1392 A 841 1 415 GSTHASGYDKTPDFILQVPVAVEGHIIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHIRPPPRAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR		1	l				
41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT 42 1392 A 841 1 415 GSTHASGYDKTPDFILQVPVAVEGHIIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHIRPPPRAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	1	1	}			i	
SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT 42 1392 A 841 1 415 GSTHASGYDKTPDFILQVPVAVEGHIIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	41	1391	A	835	7	195	
42 1392 A 841 1 415 GSTHASGYDKTPDFILQVPVAVEGHIIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHILRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR					1		SSVTNFSIGAK*DILOSVMNCLYAKRIPCVT
KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHILRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	42	1302	A	841	1	415	
QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHILRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	72	1374	΄	041	.	715	A VECUECOMITY AL TENDEMONIMIES ALLES
FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL 43		1					
KACFPTNIVTL 43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR)	ļ	}			Į	
43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR		ł]				7
LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR							l
LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
QRPMLPPSHAGLARPPPPEPISVP 44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHIRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR							
44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHILRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR		l .	ļ				
SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	144	1304	Α	852	452	1	
PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR 45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	""	1374	^	دده	472	*	l .
45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	}		1			ļ	,
45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR				1			
45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR			l				
QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR			L				
QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
WLSMSMGK 46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR			ĺ				
46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR							l ,
EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR	46	1396	Δ	900	1	366	
	10	1330	^	300	•	. 500	
VFI.FHQLNII**CLHFFIMTTFIAIPFSFLFLGR				(,		
•		L	L		L	l	VELFULLITTCLHFFIMITFIAIPESFLFLGK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosinc, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion D/KSLAMLPRLVSNSWPQVILPP
47	1397	Α	944	162	2	QLQNLASRGCL*SQLLRRLRRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK
48	1398	A	963	216	308	HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466		PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKQLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPITAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSI\RPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA/LYTWHVLEQLEILRQINQQSHGPG
56	1406	٨	1044	5	429	SVLTLQTRSPSKPLSTRKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD MH\lSLPMAFLLRTLVRCTSYIIPVTHVLSTPV TCLRRREKDGVIVDVLSDTASNHNGFPVEEH ADDTHPARLQGPTLRSQPMGPLKHKAFEERA NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFITSLIGHORMHTGERPYKCKECGKTF KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL LLLAVQQSCLADHLLTASWGGK/DPIPTKALG EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCQPFHLVMRDLLQLGQDIPQGCHY LEENHLIHRDIAARNCLLSCAAPTRAATIGDF GMARYIYRTRYYQLGDRAL/LPRK WMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER ANLMHMMKLSIKVLLQSALSLGRSLDADHA PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF GPLELVEKLCPEASDIATSVRNLPELKTAVGR GRAWLYLALMQKKLADYLKVLIDNKHLLSE FYEPEALMMEEEGMVIVGLLVGLNVLDANL\ CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE HERITDVLDQKNYVEELNRHLSCTVGDLQTK IDGLEKTNSKLQERVSAATDRICSLQEEQQQL REQNELIR
63	1413	Ā	1083		615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI HTGEKPYTCGECGKTFRQSANLYAHKKIHTG EKPYTCGDCGKTFRQSANLYAHKKIHTGIEKP YKCKECGKAFKSYYSILKHKRTHTRGMSYEG DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE KAFNHTSICCRHKKN
64	1414	A	1084	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS FFSWLTTGLTTQQRTAIE\NATVAFFLQC\\SC HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI S\GFIR\RLFLQLMLEDEKVTMFLQSPCPLYKG RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA SVALHKLSNALV
66	1416		1095	3	493	HETCSVITHIVSFSLPFLNPSHPASTPGHTENEQ PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD TLPVAAAFTETVNAYFKGADPSKCIVKITGE MVLSFPAGITRHFANNPSPAALTFRVINFSRLE HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL MTHLK
67	1417	Α	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA PYYFLLDLCCSDILRSAICFPFVFNSVKNGST WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
L	<u> </u>			sequence		nucleotide insertion_
			L			RYL
68	1418	Α	1106	1 .	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
						YEREGMQDWKTASGQSEEATQQSSQKPQPH
ł	ļ	j		1	ļ	YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
			1		}	PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
	l			١.		HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
l			ł	}		RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
	l			i	ł	SPAALAPRAARGGSRAAALAGAEAEEPLRTL
		ļ	1			APRPTRAAAPPPPPPPPPPPLPPGAPPPPVRCVSR
ļ		ļ	ļ	}	j	RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
İ]	Ì		APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
1		ł		ļ		GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
						SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
ł	ł			l		HQNIHPENQELVRVLREQLQTEQDAPAATRQ OFYTDMYCPICLHQASFPVETNCGHLFCGSLT
1	i	Ì	1			PNSIW
69	1419	A	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
07	1417	 ^	1107] 2	400	FWOHFFDLSOWKHCCEELMKIEIMSPRKPPLF
			1	ĺ		LTKEATSVYHDMSIDSPMKLESLTDIIOKKIEE
			1			TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
<u> </u>			1			RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	A	1111	698	23	ALRRLHYVRATKVVFLSFRRPFWREEHIEGGH
· ·	1			0,0		SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
İ	j	1)	j		AFAGLSREEALRLALDDVAALHGPVVRQLW
j			}			DGTGVVKRWAEDQHSQGGFVVQPPALWQT
					!	EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
		1				KSALRAAIKINSRKGPASDTASPEGHASDMEG
1		Ì	Ì			QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
İ	ł					QNTTHTRTSH
71	1421	Α	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
	ļ					PPGPPEQAGLSQFHLEPETQNPETTEEIQSS\LQ
,	ļ					QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL
						EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
	[GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI
						EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
						QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
	}					GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
	}]		HSYSICHRDLKPENLLLDEKNNIRIADFGMAS LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
						RADMWSCGVILFALLVGALPFDDDNLROLLE
						KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
	(LSLEQIQKHPWYLGGNFIS
73	1423	A	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
				-		FLGNPDKCPVQQAMLEPLGSKTETLDLRAE
	}			ŀ		MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
						SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
						LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
		1				GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
						TAGLNVAAEGARARDMPAQAWDLVERMKN
						SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
						HLATEYVQHIQQALDILSE
74	1424	Α	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
						DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
						VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
					İ	EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
						AADPAPVHTTAHPKGA .
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV
		لـــــا	L			TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ DDESGQKKLHGLQAILVHEASGTTAITATAT GYOESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVQYIMNCMLWK KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	I	1293	MAESASPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYQPCDDMDCLSDRCKILQVFDDFIFIFA MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PPYYQPEEDDEMPFICSLSGDNGIMGCHEIPP LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	i	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI CHKVIMGAGKLPRHMRTHTGEKPYMCTICE VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82	1432	Α	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	Α	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC WGRGHGCGQEALSTSHGYHLFCALLTGFLFA SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES GACVNQVTVDSITPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC NLGFFLEGHAVLTCHAGSENSATWDFPLPSC RADDACGGTLRG/AEWHHI.QPPI.PLG/ATKN

NO of mucl- ontide outlet	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuclectide seq unice cloided seq unice cloid s							
Sequence							
Sequence 91495 Orresponding the property of the property	eotide		j	USSN	location		
amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence per seq	seq-	_		09/496	correspondi	to last amino	
Presidue of peptide sequence	uence		ł	914			
Peptide Sequence			1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
NADCTWTILLELGDTTALVFIDFQLEDGYDFL	1 .		}	}	peptide		/=possible nucleotide deletion, \=possible
BTTOTEGSSLW	1 .				sequence		nucleotide insertion
87							NADCTWTILAELGDTIALVFIDFQLEDGYDFL
						_	
	87	1437	Α	1216	226	964	GTARFGPMVGFGANRRAGRLPSLVLGVLLV
	!		!	}	ł		
	}			}			
					ĺ		
BENIKKLADOPLEGOKOĒTOKIOŚNOGKELDI NNOVYPKNIPK VAENVADKNEPSSNHIPHG RVKSAGAWIHPYSDFRFYWDLIMLLIMON LIVLPVGITFFKEENSPWIPKILAWON LIVLPVGITFFKEENSPWING	•		ļ .		!		
NNQVYPKNIPKY VAENVALÄNEEPSSNHIPHG 88							
1438							
RVKSAGAWIIHPYSDFRFYWDLIMLLIMVĞN LIVLPVGITFKEENSPPWIVFNVLSDTFFLLD LVLNFRTGIVVEGAGELLAPRAIRTSYLRTW FLVDLISSIPYDYIFLVELEPRLDAEVYKTAR ALRIVEFIKLSLERL 89 1439 A 1223 1 743							
LIVLPVGITFFKEENSPPWIVPINUS.DIFFELLD LVVINFRTGIVVEEGABILLAPRAIRTRYLRTW FLYDLISSIPVDYIELVVELEPRIDAEVYKTAR ALRIVRFTKILSLLRL	88	1438	A	1218	1	534	
LVLNFRTGIVVEGGAEILLAPRAIRTRYLRTW FLVDLISSIPVDVIFLVVELEPRLDAEVYKTAR ALRIVRFTKILSLIRL 89							l l
89 1439 A 1223 1 743 MGFDEVFMINLRIRGDRRERMLRALQAQEIE CRLVEAVDGKVGMLTRSNAAPGHLAMLET LVVVAPRFVDADNILINPDTLSLILIENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGGFAPPWTSPLIDLRKAASRNL VAFYPPHDYTWSFDDIIVFAFSCKQAEVQMY VCNKEYGFLFVLRAHSTI,QDEAESFMHVQ GPWDFKFLCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLAWDSWMTGUSFTRVYLLG 1440 A 1227 2 349 NKTSPIFYLKNIVVADLIMILTPFRIVEDAGF GPWDFKFLCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLAWDSWMTGUSFTRVYLLG GPWDFKFLCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVVRHLAWDSWMTGUSFTRVYLLG 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VOLUFLVDSNPFPFGVSTVTVSTKVASMAF QTQAGAQPIPERLASERATIVKVPNNSDWAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHQLNYTLLDGRYLSEPEPPLAVVTLBSRPA PNEHINCSASRRIPESLQGADHRPYTFFISPGT RDPYGSYRLNLSSHFRASLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRASLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRASLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRAVLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLPVPSHRASLEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLAASASRAPPTPTSSAALLRERERLS TOTAVGLYSSVVVYPVYLALLFLFRNSSKV GWGWGGGGTGGGGWAAAACT GROWNERVARATOR TOTAVGLYSSVVVYPVYLALLFR HLTAFGTSLAASSAARTVETTSSSVD NFNEULESYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOEDLKQIVKH PGRKASCSTAGSGSGCLPPASSPMVSSAHPPN KAEIPERKROSTSTPNNLPSSMMTRANTYCT KAEIPERLANGTSTSTPNNLPSSMTRANTYVCT	}						,
89 1439 A 1223 I 743 MGFDEVFMINLRRRQDRERMLRALQAQEIE CRLVEAVDGKVGMLTKSNAAPGRHLAMLET LVVVAPREVDADNLLINPOTLSLIJAENKTVV APMLDSRAAYSINFUGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL LAFYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLFVPLRAHSTLQDEAESISMHVQ LEVMYPSSPSSAQSMAVVSADHIGLVISVL 90 1440 A 1227 2 349 NKTSFIPTYLKNIVVADHTTFPFRIVEDAGF GPWDFKFLCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SPTRVYLG 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPPFFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEPEPT/LAVYLHSEPR PNEHNCSASRRIPESLQADDHRPYTFSIPGT RDPVGSYRLALSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRVFPFFTADVNYIVML TCAVCLVTYMVMAILHKLDQLDASRGRAIP FCGGRGRKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSI.DIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTIFTLVADWLSVETEANGGLVEKEVL AASKASRRVPTESAALLFRRRTLVAELQRGF FDKHWLSIWDRPPSAALLFRRRTLVAELQRGF FDKHWLSIWDRPPSAALLFRRRTLVAELQRGF FDKHWLSIWDRPPSAALLFRRRTRQAATCCVLLICL FLGANAWWGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLLILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLLGRQHT 92 1442 A 1246 5 562 VFDEENLINELDDFLREEVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFOPGDYNIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOENELKQIVKH PGFRULEEYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOENELKQIVKH PGFRULEEYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOENELKQIVKH PGFRULEEYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOENELKQIVKH PGFRULEEYPMMRRAFETVAIDRLDRIGKKN SILLOKFQKDLNTGVRNOENELKQIVKH PGFRUCEYPMNRFSFTSTGEAELKEERL PGRKASCSTAGGSGRGLPPASSPMVSSAHNPN	1 1						1
89 1439 A 1223 I 743 MGFDEVFMINLRRRQDREREMILRALQAQBIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPREVDADNILLNPDTLSLLIAENKTVV APMILDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRBRRGGFAPPWNETFILDLRKAASRIN. VAFYPPHDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSTI,QDEAESFMHVQ CLEWMYPSSPSAQSMAVVSADHIGLVISYL LEWMYPSSPSAQSMAVVSADHIGLVISYL LEWMYPSSPSAQSMAVVSADHIGLVISYL VARYWEVKHLWNDSSWTGJSFTRVYLLG GPWDFKFILCRYTSVLFYAMMDTSIVVLGLIT/YDRY/WKVVRHLWNDSWMTGJSFTRVYLLG LGARLVWFGKLILAKGGHGGISWL LGASLVWFGKLILAKGGHGGISWL LGASLVAVHJANDAARG GHRSSANSVVQPQAFVGAVVTILDSSNPAAV LHLQLNTLLASSRATIVKVPINSDWAAR GHRSSANSVVQPQAFVGAVVTILDSSNPAAV LHLQLVTLLGSSPAVLHISSSPPAVLVHISSP PNEHNCSASRRIPSLQGADHRPYTFFISPGT RDFVGSYRLNLSSHFRWSALLEVSVGLYTSLC QYFSEEDVWRFGCLJLFLEETSSPQAVCLTR HLTAFGTSLFVPDSHIRVSDGPTASCHVILSSNPAAV LHQLNTLLDGLTLASGGAAP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSSSSWRITWHDNKGLSPAWFLQHIIVTD LQTARSTFFLVNDWLSVETFANGGLVEKEVL AASKASFRVPTSSAALRFRILVAELQRGF FDKHIWLSIWDRFPRSCTTRIQRATCCVLLICL FLGANAWWGAVGDSAYSTGRVSRLNPLSV DTVAVGLUSSVVVVYTALILFLFFRNSSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKOVHQRLLGKGQHT LFANADNFFVTAMSKLRFEVFQFGDVIREG AVGKMYFGHGHCVAGKKSKEMKLTDGS YFGEICLLTKGRTASVRADTYCRLYSLSVD NFNEVLEETYMMRRAFETVALDLDRIGKKN SLLAKFFLYGPGDDVIREG AVGKMYFGHGHCVAGKTSSKEMKLTDGS YFGEICLLTKGRTASVRADTYCRLYSLSVD NFNEVLEETYMMRRAFETVALDLDRIGKKN SLLLQKFQKDLNTIGVFNNQENELKQIVKH TVPPPPGGFSPAPLHPKRSPTSTGGAELKEER. PGRKASCSTAGSGSRGLPPSSPMVSSAHPN KAEIFERRKDSTSTFINNLPSSMMTRRNTYVCT							· ·
CRLVEAVDGK VGMLTRSNAAPGRHLÄMÄET LVVVAPRFVDADNLILNPDTLSLLIAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTEILDLRKAASRNL ARYPPHPDYTWSFDIUPAFSCKQUAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMYPSSPSSAQSMAVVSADHIGLVISYU LEVMYPSSPSSAQSMAVVSADHIGLVISYU 90 1440 A 1227 2 349 NKTSFIPYLKNIVVADLIMTLTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVYRHLWDSWMTGISFTRYVLLG LGARLVWFGKLILAKGGHGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRWVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAFANLSD VVQLIFLVDSNPPFFGYISNYTVSTKVASMAF QTQAGAQPIPELASFRATIVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHUQLNYTLLDGRYLSEEPPTLAVYLHSEPR PNEHNCSASRRIPESLQADHRPYTFISPFGT RDPVGSYRLNLSSHFRWSALEVSGLYTSLC QYFSEEDVWRTGLJLFLEFSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPPTADVNYIVML TCAVCLVTYMVMAAILHKLQDLDASSRGRAIP FCQQRGRKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWRRVWHDNGLSPAWFLQHHVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASRRVPTPSAALLRFRRLIVAELQRGF FDKHWLSWDRPPSSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLUSSVVVYPVALILFLFRMSRSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGYHQRLLGKQOHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDVIRGG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEIGLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLDRIGKKN SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDLBRIGKNICH SILLOKFQCDLLITGKRATASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDCHGKKN KAEIPERRKDSTSTIPNNLPSSMMTRRNTYVCT	-00	1420		1002	1	742	
LVVVAPRFVDADNI.LI.NPDTI.SLLIAENKTVV	69	1439	A.	1223	1	143	
APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL AYYPPHPDYTWSFDDIIVFAFSCKQAEVQMY VCNKEEYGFLPVPLRAHSITLQDEAESFMHVQ LEVMVPSSPSSAQAWAVSADHIGUISFU. 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMITLTPFRIVHDAGF GPWDFKFILCRYTSVLFYAMMDTSIVVLGLITI YDRYMWAVRHLWDSWMTGIJSFTRVYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESFSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAFANLSD VVQLIFLVDSNPFFFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVWQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PMEHNCSASRRIRPSQADHRPYTFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPBFMYDNYTVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATT HSLGSMWKRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFTLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWGGAVGSAYSTGRVSRILNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVMGRLGKGGVGYITSSKERMKLTDGS AVGKKMYFQHGVGTIKSSKERMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRIJGKKN SILLQKFQKDLNTOVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TYPPPFGGPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEPPERKRDSTSTPNLPPSSMMTRRNTYVCT	l i						
PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNIL VAFYPPHPYTWSFDDIIVPAFSCKQ\AEVQMY VCNKEEYGFLPVPLRAHSTI.QDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL 90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMILTPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/MDSWMTGIJSFTRVYLLG LGARLVWFGKLLAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGF)SRAPANLSD VVQLIFLVDSNFPFGYISNTVTVSTKVASMAF QTQAGAQIPIERLASERATIVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHIQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSHFWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LVGVDSRSGHRHDLOGRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFILVNDWLSVETEANGGLVEKEVL AASKASFRVPTSNAALLRFRELLVAELQRGF FDKHWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVYYVYYLALLFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPGPGDYIREG AVGKKMYFIGHGAVGVITKSSKEMKLTDGS AVGKKMYFIGHAGVGVITKSSKEMKLTDGS AVGKKMYFIGHAGVGVITKSSKEMKLTDGS VFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKLDLNTGVFNNQENEILKQIVKH PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSSMMTRRNTYVCT		•	ł				
AFYPPHPDYTWSFDDIIVFAFSCKQIAEVQMY VCNKEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL LEVMVPSSPSSAQSMAVVSADHIGLVISYL							1
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRYWKVRHL/WDSWMTGIJSFTRYYLLIG LGARLWFGKLLAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPGRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGPISRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHIQLNYTLLIGRYLSEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTGGLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHRFVFPEPTADVNTVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP PCGQRGFKYVEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDDKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRYPTPSAALLFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL PLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGFGSTGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VPDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVPGGDYIIRGG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKGIVKH 93 1443 A 1249 180 901 TVPPPPGGFSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSSMMTRRNTYVCT	[i i
LEWMYPSSPSSAQSMAVVSAÖHIGLVISYL]						
90 1440 A 1227 2 349 NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GFWDFFILCRYTSVLFYANMDTSIVVLGLITT YDRYFKPLCRYTSVLFYANMDTSIVVLGLITT YDRYFKVVAVVRHL/WDSWMTGISFTRVYLLG LGARLVWFGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVAPQAFVGAVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PPEHNCSASRRIRPESLQADDHRPYTFFISPGT RDPVGSVRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRIFLUGDRAFHRNSLDIFQATT HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLEEINVFNCRKLVATMP LFANADPINFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NINPEVLEEYPMMRRAFETVADDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSFTSTGEAELKEERL PGKKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERKKDSTSTPNNLPPSMMTRRNTYVCT							,
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILITRSRVLNEEPLTLAGFSRAPANLSD VVQLIFLVBYPFFGYISNYTYSIKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVIVQPQAFVGAVVILDSSNPAAV LHLQLNYTLDGRYLSEFPFYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVOBAYSTIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEFYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRQLPPSSPMVSSAHNPN KAEIPPERRKDSTSTPNNLPPSMMTRRNTYVCT	90	1440	Α	1227	2	240	
91 1441 A 1245 3 1937 LGSEDVRYGKLILAKGGHGGISWL 91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGPISRAPANLSD VVQLIFL/VDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVIVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEFEPYLAVYLHSEPR PNEHNCSASRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPIPSVAALLRFRRLLVAELQRGF FDKHIWLSJWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTIGRVSRLNPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGFGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPHYTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEFVAMRASPTSTGEAELKEERL PGRKASCSTAGSGSRQLPPSSPMVSSAHNPN KAEIPPERRKDSTSTPNNLPPSMMTRRNTYVCT	100	1770	^	ILLI	2	343	
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLITLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHQLUNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASSGRAIP FCGQRGRFKYEILVATGWRGGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFVPTPSVAALLRFRELLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWGGAVGDSAYSTGRVSRLNPLSV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDFLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHFKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
91 1441 A 1245 3 1937 LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLITLAGFSRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASSGRAIP FCGQRGFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKKSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPISSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1		!				1
SALTPILTRSVUNEEPLTLAGF\SRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQPIERLASERAITVKVPNNSDWAAR GHRSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEFEPYLAVYLHSEPR RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEFTADVNYIVML TCAVCLYTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCV1.LICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVALAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKGIVKH 93 1443 A 1249 180 901 TVPPPPGGPPSPAPLHPKRSFTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	91	1441	A	1245	3	1937	
VVQLIFLVDSNPFPGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVIVQPQAFVGAVVTLDSSNPAAV LHQLNYTLLDGRYLSEFPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIPQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT				.55		2507	·
QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVILDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	j j						
GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHIQLNYTLLDGRYLSEPPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRPVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FGGQRGFKYELVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSNAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCTTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSGGLPPNSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGABP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQLATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1		ı				
RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFK YEIL VKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1						
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPSVAALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPPSSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT						ļ	
FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT						·	
HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPYVLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			i			l	FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM
LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1 1						
AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLFFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			1				HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD
FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSIGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	1		}			ł	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
FLGANAVWYGAVGDSAYSIĞRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVIAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT) l						AASKASFRVPTPS\AALLRFRRLLVAELQRGF
DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPFGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	j						
GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT 92 1442 A 1246 5 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT] [ſ		!	FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	[1				-
92 1442 A 1246 5 562 VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH 93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	92	1442	Α	1246	5	562	
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT]						1
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	į		1				•
93 1443 A 1249 180 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT			ŀ				
93 I443 A 1249 I80 901 TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT						ŀ	NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT							
KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT	93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS	}	-	I	l	l	}	
	L						ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP
94	1444	A	1261	3	385	RQEDHLSPGGRGCSEL KFSQWGLTKPKLSNASP/WISLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR
95	1445	A	1282	2	550	GPRDNPGIEDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILKLARH STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENIEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA
	1446	Α		1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRKQKEQSGGEEKDED EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97	1447	A	1295		2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
98	1448	Ā	1304	118	453	SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
99	1449	Α	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP TTCTVLLLAESEGERERWLQVLGELQRLLLD ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ QLTLSPSAGLLVVLCGRGPSVRLFALAELENI EVEVPKIPESRGCQVLAAGSILQARTPVLCVA VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL GAGLVPEELPPSRGGLGEALGAVELSLSEFLL LFTTAGIYVDGAGRKSRGHELLWPAAPMGW GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL KKVRPLNPEGSLFLYGTEKVRLTYLRNQLAE KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE EQQKQQRREMLKDPFVRSKLISPPTNFNHLV
100	1450	A	1318	918	190	HVGPANGRPGARDKSP SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA LNDSMINETARDAAR VQVASTLSVLVGLFQV GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\ PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPERMYRRTVR SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN \DSCLKQKARRLTILLL
103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQVC SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	[residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	[[peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
	 	 		sequence	 	FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
		l	į	1		FTYNVGALGGALAPIIGALIAQRLDLGTALAS
	l	1	}	1	į	LSFSLTFVVILRNRRPGKSLVR
109	1459	A	1402	15	387	VLVALPDT\VTSETVVTEVLGHRVTLPCLYSS
			1	~~	***	WSHNSNSMCWGKDQCPYSGCKEALIRTDGM
		ļ		}	j	RVTSRKSAKYRLQGTIPRGDVSLTILNPSESDS
						GVYCCRIEVPGWFNDVKINVRLNLQRASTT
110	1460	A	1421	3	350	HEDLSSLLTRGSGNQERERQLKKLISLRDWM
	İ					LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
	ļ	1				FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
l		1		ł	1	CLLLASSPFPLFILLASL
111	1461	A	1426	2	344	FTSTMTKPFEKESEQPA*ATLAFGAQTSTTAD
		l				QCALKPDLSYLN:\\SSSSSSTPATSAGGGIFGSS
		l				TSSSNPPVATFVFGQSSDPVSSYGFVNTAESST
						SDSLLFSQDSKLATTS
112	1462	Α	1434	46	372	TTSWTTSCTRSCT*SGASSGPGWTPRTTWWR
		1				SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC
		ŀ	1			STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
	 	 		L		MSSSTTSSTTSTF
113	1463	A	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
i	ł	Ì				MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR
114	1464		1463	ļ ,	396 .	GIFLTFLLSNFLIVCVLLFYVSFYLFQSCINFVL
114	1404	Α	1403	1	390 .	KQQAVPEPHSSTTTPQEQEQNWYGQDLLNLQ
		1	ļ			QRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP ATSQGP*SPFSEEPPLPPSNEEVPPTLPP*EPOS
	}	ł	1			EDP*KNA*LKQMHAATTHWQQHQQHQVGC
						OYHGIMO
115	1465	A	1464	291	2	AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG
					_	GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
		1				MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
						NYCN
116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ
						YWTKYQVWEWLQHFLDTNQLDANCIPFQEF
[1			DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ
						HLKWNGDSLFLCLSLPC
117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
						VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
		j				QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN
		ł				NPEEELASDPNNEESL*RPWALEDFEIGRPLG
118	1468	A -	1485	3	385	KGK TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
110	1700	1^	1403	ا	202	PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL
						QHPWIEGHTCLDNNIHQAASEPINNNFAESKR
			1			NLAFLATGVVRHMRKLFMGANLEGPGPTVS
		ľ	ĺ			H
119	1469	A	1486	1	398	GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL
						NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE
1		i	i i			KALTKFLKWVNWDLPQEAKQALELLGKWK
		1		İ		PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
 		1	}			DDEDLLMYL
120	1470	Α	1497	3	999	MGESPAV*GYFVLAGMNSAGLSFGGGAGKY
			}			LAEWMVHGYPSENVWELDLKRFGALQSSRT
		١ .				FLRHRVMEVMPLMYDLKVPHWDFQTGRQL
						RTSPLYDRLDAQGARWMEKHGFERPKYFVP
1		([. 1	PDKDLLALEQSKTFYKPDWFDIVESEVKCCK
						EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS
						NDLDVPVGHIVHTGMLNEGGGYENDCSIARL
						NKRSFFMISPTDQQVHCWAWLKKHMPKDSN
L			L	L		LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide seq-	peptide seq- uence		in USSN 09/496	nucleotide location correspondi	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	İ		peptide	ĺ	/=possible nucleotide deletion, \=possible
		<u> </u>	ļ	sequence		nucleotide insertion MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
						GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	Α	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
						WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE MLVAHTHTVEEHTGTHLOYVSWPDHSVPDD
						SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT WDPGTDTALGWSKOPSOSYTLFES*VGSGYII
						DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
						RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
j						HVAADRG
124	1474	Α	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
						HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
						HLYEERGGVIDITAWDKDAGKRDDFIGRCOV
						DLSALSREQTHKLELQLEEGEGHLVLLVTLT
		į			ļ	ASATVSISDLSVNSLEDQKEREEILKRYSPLRI FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
					Ì	PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
						*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
						ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY
						KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK
						LELQLEEGEGHLVLLVTLTASATVSISDLSVN
						SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
			•			THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
						CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQOT
						TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
						KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	Α	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
						VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
						FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA
						WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
100	1.400		1216	-	405	WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM
						TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
						MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA LLCVWALSLVIYIGPLLGWRHPAPEDETICQI
						NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
100	1.00		1.00			AKTE
130	1480	Α	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
						EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP
						KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT
131	1481	A	1651	607	3	KDKDIK*LLFNLYSSVEILPEVLHLKT LAEGGDVFDCVLNGGPLPESRAKALFROMVE
13.	1701	A	1031	307	٦	AIRYCHGCGVAHRDLKCENALLQGFNLKLTD
						FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ
لـــــــــــــــــــــــــــــــــــــ						GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TDIPKMLWQQKGVSFPTHLSISADCQDLLK RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
132	1482	A	1656	150	48	LSNKVGGESKPKKKK LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV VDAQ
133	1483	Α	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKERQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSSPRDRDRER*RIREREREDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	Α	1678	525		ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
	1487	A	1680			AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSQKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHPSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST PEEFSRRLKERMHHNIPHRFNVGLNMKATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRGQQGWDRK YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence .		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Į	ŀ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	344	/=possible nucleotide deletion, \=possible
I	Ì	ļ		sequence	}	nucleotide insertion
					 	YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
i						NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
i		ŀ		1	•	ISSGAIYLASSYODKLRVICCKGNLVKESGTE
ľ		ĺ	{		i	HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	Ā	1686	2	526	GRPOGPAPGAGSPPESGPGLWAALGCSLVWV
1			1000	_	520	PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
ŀ	•				Ì	PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
l	l	İ	1		}	CWTRGCOTTARTAAAAAAAPGPAGRRPPGGA
	*	l				PONGSCAASASQEAAAPPPMCPPGRRWAVAS
		}				PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
137	1407	 ^	1075	,	370	FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
1		l	[IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
		1	}			
140	1490		1704	3	376	RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	3/0	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA
l		1				HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
		ĺ	1			KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
141	140	 	1543			LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	Α	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
		ŀ	·			DKLELELVLKGSYEDTQTSFLGTASAFRFHY
j		ļ				MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
						PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	1	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
						LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
l						GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
i					,	LLQVGDRVLSINGIATEDGTMEEANQLLRDA
						ALAHKVV
143	1493	Α	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
i i						KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
						NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
						SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
						LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	ì	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
ľ						PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
						KCNGEWVSQNDHVTQEGLDEATGLRVREVH
!						IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK
						SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
						CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
[THLALCPIVQHPEDTCIHSREVGVVCSRYTDV
				İ	İ	RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
						PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
						SMAAET*HHVPASGADPYVRVYLLPERKWA
					,	CRKKTSVKRKTLEPLFDET
147	1497	Α	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
						VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
						TSVTYSMG*HGAPTGSEAGANWNH**LHAH
						YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
						0
148	1498	Α	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
					·	IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
			<u> </u>	·		GIEGRLTADQLNSATACIFAAEVAIKESERFN
						GIPALSVPVAEPIRHAEALMQQALTLKRSDET
		ĺ				RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
· .				l		VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
		11	1500	· · ·	~ ⁻₹	PSQIRVVATATLRLAVNAGDFIAKAQEILGCP
				1		VQVISGEEEARLIYQGVAHTTGGADQRLVVD
				l		-4-reamentative Authorn Air

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion IGGASTELVTGTGAQTT*LFSLSMGCVTWLER YFADRNLGQENFDAAQKAAREVLRPVADEL RYHSWKEVRGASVTVQALQEIMMAQGMDE
150	1500	A	1894	2	750	RITMEIWPVD GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS PESSILDGMIRQLQQQDQRMGADQDTIPRG LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV EGVRQMHQNAPRSQIATERDLQAWKRRVVV PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	Ā	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA MEMQIKKQFQDTCKVQTKQYKALKNHQLEV TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI NEMMASQALRLDEAQEAECQALRLQLQQEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	Α	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL RQKIFKERALPDIENYMFENHDQLRQAATEC MCNMVLHKEVQERFLADGNDRLKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS QIESDRLALLQVRRILQQLGLD
155	1505	Α .	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLLKDMDLLRGREAELKQRVEAFESYQLELK DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	A	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	Ā	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNYTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR
160	1510	Α	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	~	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	цепсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucna		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
delice	l	1	1717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
)]		sequence	/=possible nucleotide deletion, \=possible
	1			peptide		nucleotide insertion
	<u> </u>	├ ─		sequence		
ĺ		1		[.		ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
	ŀ	ļ			İ	PASDFSGALETDLKASLFDQPLSIICGDSDTLP
}	ł	ł	ł	ł	l	RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
	1		ļ	Ļ		KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
<u></u>						RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	Α	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
		i			ĺ	KFQGRWGTVCDDNFNIDHASVICRQLECGSA
i	ł		i	İ		VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
İ		i		1		CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	A	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
l	}	ŀ		Į		LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
1	į		ļ	(FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
1.0.		1	1 20.2	~07	1	SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
	l ,	l	1	1		ENNWYFVVADSSKAGFTTIYKWERETGFYSH
1			ŀ	1	}	1
165	1516		0012	2	1 403	QSFTR
103	1515	Α	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
['	Í	ĺ	[NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
		1				DTHWRVAHERDELWRAQIVATTVMLERKLP
		l	l		{	RCLWPRSGICGREYGLGDRWILRVEDRQDLN
		<u> </u>				RQRIQRYA
166	1516	Α	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
		i				QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
	l					NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
1	ľ	ĺ	ĺ			SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
[WDMRNLATIFLAVVMALLSLHCLAAFKRLE
						HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
i i	}		}			VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
			1		,	LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
1						GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
[HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA
]	1511		2023	0,00		ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
						PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
1						GRLFAVVHFASRQWKVTSEDLILIGNELDLA
						1 .
						CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
		ı	1			RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
160	1510	<u> </u>	2046	<u> </u>	266	TTPQTVLRINSIEIAPCLL
168	1518	A	2046	4	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
						RLQGAARVFMPLQAQVKAKASKPLQMQIKA
[PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
اـــــا						KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	A	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR
						TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
				<u>'</u>		EDNSRSKREGLFHENECIVKINNVDLVDKTFA
						QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
					1	VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
		i				NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
						SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
					[TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLO
						SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
						ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	A	2050	363	1	
1/0	1320	A	2030	202	*	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
						VAKVEKTYDKTLENAVVADAVASKCSVLNE
						KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
		نـــــــا				ESSSEESLGESKEQLGDDVTKPSSQKA
171	1521	A	2055	139	675	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS
						DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
						LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RESCTITAYPCDSYQDYRNGKCVSCGTSQKE SCPLLGYYADNWKDHLRGKDPPMTKAFFDT
172	1522	A	2056	3	361	AEESPFCMYHYFVDIITWNKNVR LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK MEWKVKIRSDGTRYITKRPVRDRILKERALKI KEERSGLTTDDDTMSEMKMGRYWSKEERKQ
173	1523	A	2060	1	387	HLVRGKEQRRRREFMMRIRLKCLKES GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA G
174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE RRLKEEEEARLKYEKEEMERLEIQRIEKEKW HRLEAKDLERRNEELEELYLLERCFPEAEKLK QETKLLSQWKHYIQCDGSPDPSVAQEMNT
175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV AESSTVGWLCALFRVTHVGVGATGHGVVCG RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ RLQFPYLEPGHELPATTLLAFLAAV
176	1526	A	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT GIHSVYTVYQGHEIMFHVSTMLPYSKENKQQ VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PCI
177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC CDGAWLAWACWVFGNDFPSPASAACSALLG CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLO
178	1528	A	2104	2	409	ALOSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQGGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD
179	1529	A	2111	l	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLACLG
180	1530		2116	3	366	TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV
181	1531	A	2117	2 .	386	YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI
182	1532	A	2123	1	493	RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS
183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND NPPVFTRASYRVTVPEDTPVGAELLHVEASD ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW PKNNFNGSLVQASYQHEELRREVIMLACSFG NKHCHQQASTLISDWISSNRNRIPLNVRDIVY CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI DVIHVARNPHGRDLAWKFFRDKWKILNTRI RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYIVHPLWETWA HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP LDEQNRDWQGLLENLHVELTLDEEDSEGPEK EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDQPG QHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP AVVVPYMMVLQENGYGVEEGIPTLLMAASS MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV AIVVSSLDW
190	1540	A	2179	64	399	MRLNONTLLLESFGXXRPYTSEHAPTYHQW MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	-	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR LSYVLFIQERDVHKGMFATNVTENVLNSSRV QEAIAEVAAELNPDGSAQQQSKAVNKVKKK AKRILQEMVATVSPAMIRLTGWVLLKLFNSF FWNIQIHKGQLEMVKAATETNLPLLFLPVHR SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS YTHSKGIMHRDVKPLNILCNSPRNKVILADW GLAEFYHPMRKYSVHVATRYYKSPEILLDYE YYDYSLDIWAVGVILLELLTLKLHVFEGGDN EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE LPVPMGARYIRINPQSWFDNGSICMRMEILGC PLPDPNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, 'T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY ALYVHNCRLVFLLRKDFDQADTFRPAEFHW KLDQQALAKVDGQPGKSITRQLQEMPVTIQG ISLKPS
196	1546	A	2256	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF QPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	A	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	Α	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM
201	1551	Α	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	Α	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA CLAMLLHFLDTYQGLLQEEEGAGHIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	Α	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG MGANASTSPRCWDLSSGNKKWIIQVPILASIV ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	Α	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP NASNLDKVLTDIKADKDQANDGLSSALLILY LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG
211	1561	A	2431	1	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP LMFWLHGIQLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLQLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG
212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ
213	1563	A	2445			MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AMVDGRWLDCVSHHDQLFRDEYALYRPLQV LFSVYCQLECSKLIL
215	1565	A	2464	3		GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ HGPGRHGRRVCSSQDSMADVFVHLRTAWPT CSLISGQHGPGESVSYEDDDIPAPASLLHVNA AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSCTPDLTCAVSTHATVPGVRISSC TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH STVPGVRISSCTPDLTCAVSTHSTVPGVRISSC TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSIHATVPGVRISSC TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH STVPGVRISSCTPDLTCAVSIHATVPGVRISSC TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTTDLTCAVSIHATVPGVRISSCT PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHA TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHATVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTPDLTCAVSIHATVPGVRISSRT PDLTCAVSTHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSIHATVPGVRISSCT PDLTCAVSTHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVRISSCT PDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVRISSCT PDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVRISSCT PDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVRISSCT
216	1566	A	2477	1	414	TPDLTCAVSIHSTVPGLLTSVSQTSTG FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY
217	1567	A	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS
220	1570	A	2498		1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS LEGANAEIQGLKENLQNTNALNSQTQAFIKSS

NO. of NO. of No. of N	SEQ ID	SEQ ID	Mct	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq			1	,	I .		
USSN Osarion			""				
1949 1949		, , ,	1				
1914 ng to first amino eacid residue of peptide pe	1			09/496	correspondi		
				914	ng to first	acid residue	
peptide			}		amino acid	of peptide	
					residue of	sequence	
PFNTSAEIQELRGHIENAGDEHIVLKRIDLEM			Į	1	peptide		
VITAGTOKANGLDQTDTQIQVPKSEMENVN			<u></u>		sequence		
TINAQIQVINGHMKNASREIQTILKQQMKNL							
SALTSQTOMILOSNLOKASAEIQRLRGDLENT 221							
		ļ	l	ļ		}	
1571		ļ.		}		[
TDEDTIRILSRKFKDWAYGPVYSSLYDL				<u> </u>			
TCGEASVLEIL VYNSKIENRHEMLA VEPPNE	221	1571	Α	2501	3	500	
LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT				ł	i		
AYYQPLEGTPPYPYRTTVDYLRIAGEVITLFT		ł	l			!	
222 1572							1
222		}					
ARHYDKEEEVDI.VMQKYESI.RQELRITER							
AKKELEVHTEALAAEASKORKLREQSEHYSK	222	1572	A	2508	3	395	
QLENELEGILKQKÇISYSPGVCSIEHQQEITKL		1					
		ł	ł			ļ	
223		}					
Mydertdylksi.kektppfshodqavlogs				ļ			
EASSNKDMFADRLSKSIKHSIDKSKSVIPNID KNAVYKESLPVSGESQLTPEKSPKFPDSQNQ LTHCSLSAA 1	223	1573	A	2544	2	412	
		1					
LTHCSLSAA	1		t			1	
1574							
CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI QEGLKNTIGAIIFLLMYFFGMASSIWWULTL TWELAAGLKWEAIFLHMYFFGMASSIWWULTL TWELAAGLKWEAIFLHMYFFGMASSIWWULTL TWELAAGLKWEAIFLHMYFFGMASSIWWULTL VK 1			ļ.,	2552	401		
QEGLKNTGCAIIFLLMYFFGMASSIWWVILTI.	224	1574	A	2552	401] 1	
225 1575 A 2563 724 1 MSARKERREKGEEGGEKDGDEDEKEEEKE GLGEEEKEAGKKKKQEEKEKEKGAVYSR VARICKNDMGGSQRVLEKHWTSFLKARLING SVPRGDSFFYFDVLQSITDIIQINGIPTVVQVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKKPRGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGYLFYYGNYVGDVMFEMAAEMAQEVAIP TRITVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTTRRINHFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDFIIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINKRYAFKAAHPNNRTYYFCIDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAESKYGFQKDGDRPLTKINSVKLINSL PSEYFEGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKJQKGRGHEEETRGVISYQTLPRINMFS HRAQIMAR YPEGYRTLPRNSKTPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSSPKTMVNISDQT MHSIPTSPSHGJAAVQQSYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGIGSVLPQGMWASPEMPTICSVFEA HIVILFHALTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPPIICKCKCPGYIGSNCQTALCDPDCKNH			ļ				
VK		ì				ł	1 7
1575 A 2563 724		ł	i	1		}	l e e e e e e e e e e e e e e e e e e e
GLGEEEKEAGKKKKKQEEKEKGAVYSR VARICKNDMGGSQRVLEKHWTSFLKARLINC SYPGDSFFFYDVLQSITDIQNGIPTYVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPLADEPWFT KTRVRYRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRGRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGELYTEADRVIQRTNSMQQ LEQWIKJQKGRGHEETRGVISYQTLPRIMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHALTTERGFQVTVLVGNVRQTAVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	205	1676	<u> </u>	2562	704	ļ <u>, </u>	l
VARICKNDMGGSQRVLEKHWTSFLKARLNC SVPGDSFFYFDVLQSITDIQINGIPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPLADEPWFT KTRVRYRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPPISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	223	13/3	A	2563	/24	1	
SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT QLNSIPGSAVCAFSMDDIBEVFKGRFKEQKTTP DSVWTAVPEDK VPKPRPGCCAKHGLAEA YK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRVRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKJQKGRGHEEETRGVISYQTLPRNMFS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHIDKTLGPGAEEKRRSMDDTMWQLYEW QQRQFVNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNQQTALCDPDCKNH			1			!	
QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTVARVRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRIVLTIDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKFEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAGIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIFTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGIGSVLPQGMVMASPEMMPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			l				
DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLINSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFIGIGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			ĺ	1			
TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH 226 1576 A 2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRINFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLINSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHUKKTLGFGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFIGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH							
KTRVRYRLTAISVDHSAGPYH			İ	1			
2571 449 3 EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGHGE RGVIREKMMPADAIVDHIMDRIFS		•	1	[
TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	226	1526	 	2571	440	3	
AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHIPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	220	1370	^	23/1	447	3	
VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH		[Ì		İ		
RGVIREKMMPADAIVDHIMDRIFS 227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMPTICSVFEA HIVLLFHATTFRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	1	i					
227 1577 A 2575 3 1197 VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETINIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH							
EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	227	1577	A	2575	3	1197	
ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	٠	15//	^	2513	ا	1177	
TKETNNIPNHRVLIKPĒIQNNQKNKĒMSKIĒĒ KKALĒĀĒK YGFQKDGQDRPLTKINSVKLNSL PSĒYESGSACPĀQTVHYRPINLSSSĒNKIVNVS LĀDLRGGNRPNTGPLYTĒĀDRVIQRTNSMQQ LĒQWIKJQKGRGHĒĒTRGVISYQTLPRNMPS HRĀQIMĀRYPĒGYRTLPRNSKTRPĒSICSVTP STHDKTLGPGĀĒĒKRRSMRDDTMWQLYĒW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIĀA YQGYSPQRTYRSĒVSSPI QRGDVTIDRRHRĀHIPĶVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMĀSPĒMNPTICSVFĒĀ HIVLLFHĀTTFRRĢFQVTVLVGNVRQTĀVVĒ KIHĀKVRGTWPFISPĒVRKĒGGLPQTGRĒLLD PTMĢIKPHLWWVĀĀ 229 1579 A 2589 I 448 DDKNĀQGIKRHVKPTSGNĀFTICKYPCGKSR ĒCVĀPNICKCKPGYIGSNCQTĀLCDPDCKNH		[1				
KKALEAEK YGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQK GRGHEETRGVISYQTLPRNMPS HRAQIMAR YPEGYRTLPRNSKTRPESICS VTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			1			Ì	
PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH)			
LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPVKK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH							
LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			1				
HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAK VRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			1	}			
STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH				[
QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHIPKVK 228 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH							
MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTlDRRHRAHHPKVK 228			1	i			
QRGDVTIDRRHRAHHPKVK			1	[
228 I 1578 A 2583 3 330 LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 I 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			l				
HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA 229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	228	1578	A	2583	3	330	
KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA			l		-		
229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH			}]			
229 1579 A 2589 I 448 DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH							
ECVAPNICKCKPGYIGSNCQTALCDPDCKNH	229	1579	A	2589	1	448	
			J		-	' '	,
							GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met · hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	-		914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
j .	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			l	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
020	1500	<u> </u>	0000	<u> </u>	120	ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	Α	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
						WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
		i	ſ			GRARRTPTCEPATPLCCRRDHYVNFQELGW
		j		}		RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
						FHSAVFSLLKANNPWPGRTSWCVPTARRPLS LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
255		''	200.		103	YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
			}			NSIPYWERIT
234	1584	A	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
						DGSKSSDDQKIISYLWEKTQ
235	1585	Α	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
						AKFLNVEAAMVFGMGFATNSMNIPALVGKG
			ł			CLILRDEVNHTSLVLGARLLGATIGIFKHNYA OSLEKLLRDAVIYGOPRTRRAWKKILILVEGV
		}	ĺ			YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
						GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
						FGASGGYIAGRKARILSPPACLVPNTGSHSLH
			[,		RLTRDLQMNEAMVALVTDRLQGWNSGEGN
						WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
226	1506		0(2)		200	AEQIIRSLKLIMGLDGTTQ
236	1586	Α	2621	1	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
						WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
						KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
						A
237	1587	Α	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
						GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
						LMHRLMLDCWQKDPGERPRFSQIHSILSKMV ODPEPPNV
238	1588	Α	2631	1	1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
						ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
i						WQPCSRTCGGGVQKREVLCKQRMADGSFLE
						LPETFCSASKPACQQACKKDDCPSEWLLSDW
						TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
						TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
						VVLRCPARRVRKPLITWEKDGQHLISSTHVT
						VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
						VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
	1					KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
<u> </u>	1600		0.00		(50)	RYDDLVSRLLEQGAPCSSSKKKN
239	1589	Α	2636	1	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
						TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
						TVVTYPSAWSOEMVSLLKKLLEPNPDORFSO
						LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
						GRLNCDPTFELEEMILESKPLHKKKKRLAKK
						EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE
0.00	1.506					KVNRDCI
240	1590	Α	2639	389	3	ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
					}	LWQNFAREIEEHVFTLYSKNIKKYKTCIRSKV ANI.KNPRNSHLQQNLJ.SGTTSPREFAEMTVM
		1	L			WANTA MARTINA CHARTON TOLKEL VENT AM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ 1D NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1591	Α	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	Α	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1	1271	MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599 ·	A	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	Α	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, I=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QKRKKKAPDHSSGRKEELVITHTVDKLETKK
						PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI GAGDCL
252	1602	A	2697	421		PQKSHSGAYQCFATRKAQTAQDFAIIALEDG TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT VTWALDDEPIVRDGSHRTNQYTMSDG1TISH MNVTGPQIRDGGVYRCTARNLVGSAEYQARI NVRGPPSIRAMRNIT
253	1603	Α	2698	65	401	ACCQWRRTLIPAKSTTVSCTISTPHHPFRGSYS FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG AAFTNNIASSTIIL
254	1604	Α	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA DKETLENMMQRHEEEAHEKGKILSEQKAMIN AMDSKIRSLEQRIVELSEANKLAANSSLFTQR NMKAQEEMISELRQQKFYLETQAGKLEAQN RKLEEQLEKISHQDHSDKNRLLELETRLREVS LEHEEQKLELKRQLTELQLSLQERESQLTALQ AARAALESQLRQAKTELEETTAEAEEEIQALT VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	Α	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ PYEAARMFFEGLR
257	1607	A	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD FNPSFSFLDPRYSVGGDENIGTVTTLANILREF NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL
258	1608		2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI MTRKELLTVYSSEDGSEEFETIVLKALVKACG SSEASAYLDELRLAVAWNRVDIAQSELFRGDI QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH SAGTKAPALKGGAAELRPPDVGHVLRMLLG KMCAPRYPSGGAWDPHPGQGFGESMYLLSD KATSPLSLDAGLGQAPWSDLLLWALLLNRA QMAMYFWEMGSNAVSSALGACLLLRVMAR LEPDAEEAARRKDLAFKFEGMGVDLFGECYR SSEVRAARLLLRRCPLWGDATCLQLAMQAD ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR GGFSQREMVTGERSPSPEEEEEEEEGFGERA SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN IATVVKGAERASSMAGTKPYMAPEVFQVYM DRGPGYSYPVDWWSLGITAYELLRGWRPYEI HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR EIYLQGCFKPLVSISPNDSLFFAVYTLIKNRIH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
[[peptide	ĺ	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1		ĺ			ĺ	RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
						RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
1		l	l	1	}	DIFVDRRVSALAVVNECGTHPQDERLGLGW
262	1610	 	0000			GLGEPGSEERLFPAAITSR
262	1612	Α	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
1					1	GRLVKLSLANNNLVGVHEDAFETLESLQVLE
}					j	LNDNNLRSLSVAALAALPALRSLRLDGNPWL
]		ļ	CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM ESRRISLRACRRPASRV
263	1613	A	2736	2	343	
203	1013	^	2/30	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
i l						ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
						RLISPLVNLPOSPGGLEFOYOAT
264	1614	A	2738	2	245	RAMLKCLREGOPPPSYNWTRLDGPLPSGVRV
201	1014	Α.	2750	_	243	DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
			ľ			DTVDVLDPPEDSGKQVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
1				ļ -	300	LAAVETTVLVLIFAVSLLGNVCALVLVARRR
1				•		RRGATACLVLNLFCADLLFISAIPLVLAVRWT
]] .			EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
						SLER
266	1616	Α	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
1 1						LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
						V
267	1617	Α	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
[[HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
						LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
						AVLLLLLLSLALGLVLAALGLFVHHRDSPL
1 1						VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
						ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
269	1619		2772	3	243	LPLSWAE
209	1019	A	2112	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
						LSKNLSFSEFCFDVSY
270	1620	A	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
	.020	11	2.07	•		VEQIAKAEETHSSLSQELQARLQTVTREKEEL
1					İ	LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
						KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
[i	KAYDELRLQSEAFKKHSLDLLSKERELNGKL
						RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
		ĺ				FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
<u> </u>				-	Į	RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
į i	ł	1	1		ľ	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
			l			VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
L						SQRDDGKGALNIVEFLRKREELHQILSTVKQP
272	1622	Α	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		j	.]		j	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
				l		GYRKVVSNNCTDGVREQYTAKPQKCPGKAP
]]				l		RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
	1	İ	}	ſ		VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
L						YQNXGIXRXTVQVDNSLGS
273	1623	Α	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
						DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
	1	1	j	ĺ		KADSLNVSRNSVMQELSELEKQIQVIRQELQL
774	1624		2006	160	220	AVSRKTELEEYH
274	1624	A _	2805	168	320	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE
L		1				IFIARNGVVGETLTHCKRV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1200	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
) denot			1 71 1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ł		peptide	Sequence	/=possible nucleotide deletion, \=possible
1		ł	1	sequence		nucleotide insertion
275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
2/3	1025	Α	2012	208	321	MGKIIFQ
276	1626	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
270	1020	Δ .	2013	(4)	200	
		ŀ				KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY OVGPVRRNGEAGPG
277	1627	A	2817	3	410	
} 2//	1027	A	2017	,	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
1.			i	İ	i	LFISYLHTPKHKQHEVLQAMGSILGITGEEME
1	1	ł	1		[PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
	1	1				GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
050	1600	<u> </u>	1		1.52	LPPHNSPGKIK
278	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
					ł	VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
		<u> </u>	ļ			PVDQNPRLV
279	1629	Α	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
	1					TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
ŀ			i .		1	CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
			L			VPSQRHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKQKRLALCVFSFEEGGRRK
ľ	İ	i				CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG
						VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASQEGGDVLGAR
}		l	1			QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA
			ĺ			NTTNMDEVPRPQALSGSSVVWVSGCVASRS
L						VILSLTSG
282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
			1			TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
					*	YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
)	}			YLKTLPPYYL
283	1633	Α	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
	ĺ	ł	ł			MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
			l			PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
	,	1				SSEESAP
284	1634	Α	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
						DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
						KSLAETVLNFPLDKSLLLRCSNWDAETLTED
1		!				QVTYAARDAQISVALFLHLLGYPFSRNSPGEK
						KR
285	1635	Α	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
						VCDRVSEDGINRQQAQEWCIKHGFELVELSP
		}				EELPEEDGKCLCVRRKYGTYI
286	1636	A	2845	197	278	TAEDVLTVAYEHGVNLFDTAEVYAAGK
287	1637	A	2851	2	427	FVAEVRREWAKYMEVHEKASFTNSELHRAM
						NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
i .			1			AVLONLKRILAKVOEMRDORVSLEOOLRELI
					İ	QKDDITGSLVTTDHSQMKKLFEFQLKKYDQL
		١.	1			KVYLEQNLAAQDRVLCALT
288	1638	Ā	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIOSLELDK
						LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
		1				TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
])			KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
]						EPGQELAETALHLAVRTADQTSLHLVE
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
207	1037	^	2001	-	7.74	DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
			{ [İ	
			1			DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
j l			[
290	1640	A	2868	1	378	ITDFEGQPPTPRDKLSCWVYKDRLIYFG
***	1040	Λ	2000	•	3/0	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
			[1		PDCASCLQAQDPLCGWCVLQGRCTRKGQCG
		L				ו שלאיטרועאלטו דרים אר ארלטער ועאלטלרם

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence	}	09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	"		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l	i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide sequence	[/=possible nucleotide deletion, \=possible nucleotide insertion
			 	sequence		RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR
						QE
291	1641	A	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
		Ì		Į		PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
		[GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL AKVINAENAAHKSEKFRAMATRTRQEYLKD
	}	ļ		J		LA
292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
						PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	1	427	REKEEEVEEEEDKVVKETEKEAEQEKEEDSL
].]				GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
						LSPEKLTAENRYYCESCASLQDAEKVVELSQ GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
						LRLPLAGGRGQAYDL
294	1644	A	2879	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI
						IIVFVTGGVLG
295	1645	Α	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
1						NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
						LSESISO
296	1646	Α	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
						RLQEFSQKMDQVRGHWPVST
297	1647	Α	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
						KLYSTMGRFLRDRKNPACREMAVVLLANLA
						QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
		}				LLALAKVDDNHSEF
298	1648	Α	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
						SGLLNASAQVNL
299	1649	Α	2898	1	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL GYFQAYNVLILTMQASLPKVLRFCACAGMIY
						LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
1 1						VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI
						SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
200	1650		2001		415	LQEF
300	1650	A	2901	. 1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
1						SLVTTITGSSLRGEPMSYYIASGNLGNTFOIDO
						LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
						FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
						EPCGWMYDHAKWLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	Α	2909	2	412	GPOMLCKKIYFIWVTRSOCOFEWLADIMOEV
				_		EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
						CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
			1			SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
303	1653	Α	2914	291	453	LVNRQDRAHFM KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
203	2000	n	+1رع	271	<i>U.</i>	VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
L						NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	Α	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
						DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
						PEIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP
306	1656	Ā	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
] }						EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA
L						LTKGALWAVFLLAGSALLCAEVTGVIWRQPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SKTKLSFKVSSSA
307	1657	A .	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT CTAENAVGRARRRVHLTILVLPVFTTLPGDRS LRLGDRLWLR
308	1658	Α	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM DSSLPEEEEDEDKEAINGSGNAENRERHSESS DWMKTVPSYNQTNSSMDFRNYMMRDETLEP LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP RLCKKAKAPEDC
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIDTFVGHDVVEPGSYVQMFPYPC YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI TGFVQLSISVTALTAILKYGQVLMHSHVVIIW LFLAVYAVATIMFCF
311	1661	A	2963	3	465	MKPQMPGLGAPNGYGPGRGRAGVPGGPERR PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK PQKPGLRGTLKPQKSGHGHENGPWPGPCNA RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ KI
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ
314	1664	Α .	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM
315	1665	Ā	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE
317	1667	A	2981	3		VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARI.I.RIDIANTI.REQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion STLALSHSAQVLASASGRSSTTAHCQIRVWD
320	1670	Ā	3000	693	322	VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL GDHDGRTLALWGTGHL IDESTGLIITVNYLDYETKTSYMMNVSATDQA
320	1070	A	3000	000	322	PPFNQGFCSVYITLLNELDEAVQFSNASYEAA ILENLALGTEIVRVQAYSIDNLNQTTYRFDAY TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF LFPWLFLQVEVIKKAYMQGEVEFEDGENGK DGAASPRNVGHNIYILAHQLARH
323	1673	Ā	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND ERVFGKRGF
324	1674	A	3020	523	797 .	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ EGDLVEVVLSASATFEDFQIRPHALTVHSYRA PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK RC
328	1678	A	3030	13	569	TTRPTISCQRPGPGLAAGMLPYTVNFKVSART LTGALNAHNKAAVDWGWQGLIAYGCHSLV VVIDSITAQTLQVLEKHKADVVKVKWAREN YHHNIGSPYCLRLASADVNGKIIVWDVAAGV AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI HPPNYIVLWNADTGTKLWKKSYADNILSFSF D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL HRMAEKVGADITVLREREVDYDSDMPRKITE VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL LGVLTQGELDNGRGRARLNLFRHLHEIQSGR TSSISFEILGFNSKGEVHGINGTQWGQTLRMG W
330	1680	Α.	3040	3	397	LCSTLLLTIPSWVLSQITLKESGPTLMKPTET LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ VVLTMTNMGPVDTATYYCAQFARGARGSN WFDPWGQ
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK GENRKTLISGMIDEPHAIVVDPLRGTMYWSD WGNHPKIETAAMDGTLRETLVQDNIQWPTG LAVDYHNERLYWADAKLSVIGSIRLNGTDPI VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG KRLDNGTCVPVPSPTPPDAPRPGTCNLQCFN GGSCFLNARRQPKCRCQPRYTGDKCELDQC

000 10	CCOID	1 3.7-4	LCCO	I D 3:-4-3	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	поа	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide		ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	seq- uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq- uence	uence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ucaice			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		}	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
	}	j]	sequence		nucleotide insertion
		ļ	ļ	sequence	 	WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
		1	1			TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
			į			FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
' I		İ	ł	ì	1	CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS
			Ì		}	GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT
		ļ	ŀ	İ	Ì	MNSKMMPECQCPPHMTGPRCEEHVFSQQQP
		}	}			GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG
332	1082	^	3043	*	932	AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT
		l				LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD
		i		ł	1	MPKSPFKRRSMNEIKNLQYLPRTSEPREVLF
		1			1	EDRTRAHADHVGQGFDWQSTAAVGVLKAV OFGEWSDOPRITKDVICFHAEDFTDVVORLO
		1				LDLHEPPVSQCVQWVDEAKLNQMRREGIRY
		}	}	ļ	}	
		l				ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW HIRLKOYHPVVEATONTESNSNMDCGLTGKR
		l			-	
		ļ]			ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP
222	1.602		2046	407	1/2	PASE
333	1683	A	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL
						YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT
		}				ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG
224	1704	<u> </u>	2052		076	CSSKTWKVAPFVRAWWRP
334	1684	Α	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD
ì						EASGANDEIVQLRSEVDHLRREITEREMQLTS
-025	1605	ļ.,	2054		0.46	QKQVRRVNKVVRSLEDF
335	1685	Α	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR
]]	}		NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA
		ŀ			•	YNDVQYQGHYYEWLPRYNDPAAPCALKCH
į						AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC
						RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
	•				ĺ	TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV
		ŀ			1	ENTTVEFORGSEROTFKIPGPLMADFIFKTRY
		ŀ			!	TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT
1		ļ				CGGG
336	1686	A	3058	54	347	VVGKOEAGAHSDSCCLLHTPPRLTPAHSRKA
330	1000	^	3036	54	347	LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG
		1				AWDWRLGSPACPHWGLHKLPRLWDPLSLYP
		İ	!		1	
337	1687	A	3059	2	709	VLCWGT ILTSLVELTRFETLTPRFSATVPPCWVEVQQE
331	100/	A	30J 9		109	
						QQQRRHPQHLHQQHHGDAAQHTRTWKLQT
j					1	DSNSWDEHVFELVLPKACMVGHVDFKFVLN
						SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF
						LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
						SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK
220	1600		2060	95	204	VAAGKEKSSNVKNENTSGTRK
338	1688	Α	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK
}					}	DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL
		1				SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE
	1600		2062		260	EELNP
339	1689	Α	3063	236	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV
						PSSVTTMLSWV
340		–	3065	3	1249	DLWQFTPLHEAASKNRVEVCSLLLSYGADPT
	1690	Α	2002	1 -		
	1690	A	3003			LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL
	1690		3003			QAAREADVTRIKKHLSLEMVNFKHPQTHETA
	1690		3003			QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE
	1690		3003			QAAREADVTRIKKHLSLEMVNFKHPQTHETA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IISLQGFTALQMGNENVQQLLQEGISLGNSEA DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE GRQSTPLHFAAGYNRVSVVEYLLQHGADVH AKDKGGLVPLHNACSYGHYEVAELLVKHGA VVNVADLWKFTPLHEAAAKGKYEICKLLLQ
341	1691	A	3070	1	547	HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR GDAALLDAAKKGCLARVKKLSSPDNVNCRD TQGRHSTPLHLAGK GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI RTVKFLRSATIPVVELMDVQGERLDMEVGFD NRQAAFDMVCTMLEKRVRHKILYLGSKDDT
342	1692	A	3073	463	3	RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI HLRMQLMRDALSANPDLDGVFCTN RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE IDDNRVTAEEVDILLREGEKLAPVMAKTRILR AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE RDGLDGFITITGGKLMTYRLMAEWATDAVC RKLGNTRPCTTADLALPGSOEPAKVP
343	1693	A	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS LGASRAQVLWFVILPGALPEILTGLRIGLGVG WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV AHSKPSTRNILLLL
345	1695	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC VINQAAATGGDARQLVGYLVSQSGLPLDTSA LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL DRKALPLPELKAQAPGRAPKAGSETIIAAAFS SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	A	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH
349	1699	A	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR SALTYEPIFLILLEALRGVIHATPCRVSLSLWD GLDLP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER VAGLHFFNPAPVMKLVEVVSGLATAAEVVE QLCELTLSWGKQPVRCHSTPGFIVNRVARPY YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFIILNGTFLNIGETDTESCVNGWVYDRSS FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL ESRICVVGENGAGKSTMLKLLLGDLAPVRGI RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL LGTQVFLGRPEEEYRHQLGFGMGISGELGHA SSLPACLGGQKEAEVAFCSDGLLPCPNFLNIL DEPTNHLGHGRAIEALGPCLQTISGVGVILVS HE*SALSRLVCRELLWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG AASREHARWQGTGLAPGTRVAVAPTCVQGL PQERSVCRPFFSSRWREGPVWALGAGAHGKP RWSGGVRCVVRGGRWFTPAPH
357	1707		3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ PGLYFGGAAAVAEPDHLREAGITAVLTVDSE EPSFKAGPGVEDLWRLFVPALDKPETDLLSH LDRCVAFIGQARAEGRAVLVHCHAGVSRSV AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ KVTEKYPELQNLPQELFAVDPTIVSQGLKDE VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	Α .	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP PTANREINPGPAAAADTRSCWGHKRSWRGW RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG KGAGGKPSETLTRSPPVWRGKRGSANGFLSW VQILQ
359	1709	A	3132	3	191	HEHLLLLLLCVFLVKSQGVNDNEEGFFSARG HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA *AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL *RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL OA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ APGLPHRTSIRPGWRRLTEPEAWARRHRRPW GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR QG*PSGAPW*LPGLAQLAFQCHI.PHDEVGPP

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RNQSPLGNDTLSSGLPMGPRRQVWPLARVG GHSSPREPQVLKKPLWGQTDIAGVGSASLYP DNL
362	1712	A	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI SGASLCPVLSPPR*PLOPSSL
365	1715	A .	3145	122	413	LLPYPSLFVFLRQCHFVTRLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK GLMATIYEELLKLSNKLIQ
367		A			2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPI.HSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTAIATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLQDQLRDAQQQVKALGTERTTLEGHLAKV
						QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVP MSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	Α	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*I*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	A	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410		RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSENPPI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	!	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
dence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	!		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	sequence	/=possible nucleotide deletion, \=possible
	1	[sequence	ĺ	nucleotide insertion
		<u> </u>		Sequence		
	ļ	١ . ا			 •	FEIR/MYDAQHQQVGSNKCRVNNAGCSSLCL
ŀ		[[ATPGSRQCACAEDQVLDADGVTCLANPSYVP
						PPQCQPGEFACANSRCIQERWKCDGDNDCLD
						NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
		[!		LCDGDNDCGNSEDESNATCSARTCPPNQFSC
		1				ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
		!				CFPLTQFTCNNGRCININWRCDNDNDCGDNS
!						DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
l						NDCGDYSDETHANCTNQATRPPGGCHTDEF
	•	1				QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
						GVTHVCDPSVKFGCKDSARCISKAWVCDGD
		i I				NDCEDNSDEENCESLACRPPSHPCANNTSVC
1						LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
[GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
1						CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
ļ	ĺ					WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
		1				KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
		ļ.				VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
		1				LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
		[[TLLAGDIEHPRAIALDPROGILFWTDWDASLP
						RIEAASMSGAGRRTVHRETGSGGWPNGLTV
}						DYLEKRILWIDARSDAIYSARYDGSGHMEVL
						RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
ļ						KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
						QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
						ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR
1						GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
						QRVYWSDVRTQAIKRAFINGTGVETVVSADL
		1	ĺ		. [PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
1		i				RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
						WTDGDNISMANMDGSNRTLLFSGQKGPVGL
1			1			AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
						AMRSQLGKATALAIMGDKLWWADQVSEKM
						GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
!	•					QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
i						MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
]						GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
[]			Í			TIYWVDMGLSTISRAKRDQTWREDVVTNGIG
[l				RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
						SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
			[į	ĺ	GOYPRIERSRLDGTERVVLVNVSISWPNGISV
[DYQDGKLYWCDARTDKIERIDLETGENREVV
į l						LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
[[- 1		į	
						RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
		·				QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI
		1	ľ	l		
			ļ			HLSDERNLNAPVQPFEDPEHMKNVIALAFDY
			-			RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT
			Ì		ľ	IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
			ļ		ļ	RHTVDQTRPGAFERETVITMSGDDHPRAFVL
			İ			DECQNLMFWTNWNEQHPSIMRAALSGANVL
		ĺ	1			TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
						RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
	ļ		ŀ			WTDWVRRAVQRANKHVGSNMKLLRVDIPQ
		[İ	1		QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
		i	[l		LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR
		.	1			AQDEFECANGECINFSLTCDGVPHCKDKSDE
		1	ĺ	ľ		KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
		ļ	Į	-		GADDCGDGSDEIPCNKTACGVGEFRCRDGTC
				l		IGNSSRCNQFVDCEDASDEMNCSATDCSSYF
					 	

NO: of moti-	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence unde seq- uence unde seq- uence unde seq- uence unde seq- uence unde seq- uence unde seq- uence unde seq- uence unde seq-							
USSN Ocarison 1914 1914 1914 1914 1914 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 1914 1915 191							
uence 99496 Orrespondi ng to first an ing to first and ing to first and in each residue of peptide sequence Q=Glutamine, R=Arginine, S=Serine, 1—Threenine, V=Valine, W=Tryptophan, y=typeriode sequence Q=Glutamine, R=Arginine, S=Serine, 1—Threenine, V=Valine, W=Tryptophan, y=typeriode sequence Q=Glutamine, R=Arginine, S=Serine, 1—Threenine, V=Valine, W=Tryptophan, y=typeriode, x=typ				USSN			
amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence per control of peptide sequence per per per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide sequence per control of peptide seque	seq-	uence		09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence V=Tyrosine, X=Uhxhown, *=Siop codon, /-possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion RiGVKQVI=P(PCERTSLCVAPSWVCDGAND CGDVSDERDCPQVKRPRCPLNYFACPSGRCIP MSWTCDKEDDCEHIEGDETHCHNKFCSAQPE CONJRICISKQWLCDGSDDCGDGSDEAAHCE GKTCGPSSTSCPGTHVCYPSRWLCDGINDCA DGADESLAAGCLYNSTCDDREFMCQNRQCIP KIFVCDIBRDCAADSBESPECEPYTCGPSEE RCANGRCLSSRQWECDGENDCEDGOSDEAAHCE GODDCGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGDSSDERGCHINECLSRKLSGCSQDC GDDCGGCSSCLCSGRCVARDVDGCS HICKACYDCEFFICKARDGARTAVYLKKLLDGSNY TLLKQGLNNAVALDFDYREQMIYWTDVTTC GSMIRRMHLNGSNVQVLHSTGLSSPPGCLAV USGLREPRALVVDVOMYLHTGLSSPPGLAV DWVGGNLYWCDKGRDTEVSKLNGAYRTVL VSGLREPRALVVDVOMYLHTGLSSPPGLAV USGLREPRALVVDVOMYLHTGLSSPPBCLAV USGLREPRALVVDVOMYLHTGLSSPPGLAV USGLREPRALVVDVOMYLHTGLSSPPHOLYD GSMIRRMHLNGSNVQMGTLYWTDWFINECK KIRLLIFERDLYVVTDWFTKSINAAHKTTGTN KILLISTLIRPMLLHYHALRQPDVPNIPCK VSSLREPRALVVDVOMGTLYSOMPHOLYD GSMIRRMHLNGSNVQMGTLYWTDGDHNPCK VNNGCSNLCLLSPGGGHKACPTNYLGSD GRICVSNCTASQVCKONKCPPWWCDPHNPCK VNNGCSNLCLLSPGGGHKACPTNYLGSD GRICVSNCTASQVCKONKCPPWWCDPHNPCK VNNGCSNLCLLSPGGGHKACPTNYLGSD GRICVSNCTASQVCKONKCCPPWWCDPHOLYDOCG GRODGDSDEAGEPCCPECTOCKTPAW CDDCGDSDBACPGFCCSGCCCPARW CDCGDNOSDBEACCGTARW CDGDNOSDBEACGTGNACOGODHOCADG GRODGCGGGGEBCCCGPRWWCDGDHOCADG GRODGTGRTCPLDFCYPORC KNNRCVPGRWQCDYNDCGDNSDESCCTPR CNSERSCNARGCGGGGGGGGGGEBCCCSBPCCPW CNSERSCNARGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
peptide sequence succession in section in se					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Incledité instrition					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
REGYKGVLEPPCERTSLCYAFSWYCDGAND CODYSERDCPGVKRPCELNYRACPSGRCIP MSWTCDKEDDCHGEDETHCNK7CSEAQPE CONING CISK QWL.CDGSDDCGMOSDEAAHCE GKTCGFSSFSCPGTHVCYPERWL.CDGDKDCA DGADESLAAGCLYNSTCDDREFMCGNROCIP KHFVCDHDRDCADGSDESPECKPYTCGFSEE RCANGRCISSQWLCDGGBNDCHOSDEAPK NPHCTSPEHKCNASSOFLCSSGRCVAEALLCN GGDDCGDSSDERGCHINECISKLSGCSDC EDLX(GFXCRCRPGFRLKDDGRTCADVDECS TTPFCSGRCINHGSVKLCLVGGVAPRGGIP HSCKAVTDEEPFLJFANRYYLKILALDGSNY TLAKQCLNNAVALDFDYREGMYYDTVTTO GSMIRMHILMGSNVQVLHRTGLSNPDGLAV DWYGGNLWCDKGRGTTEVSKLNAGAYRTYL VSSGLBEPRALVOVDNOYLYNYTDWGDHSL SIGRIGMGSSSSVIPGTSWCRAGAYRTYL VSSGLBEPRALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNYTDWGDHSL VSSGLBERALVOVDNOYLYNTAGAGCCGTUNGGCGGTGCGNAGAGGNAGAGGNAGAGGNAGAGGNAGAGGNAGAGGNAGAGGNAGAGGNAGAGGNA							/=possible nucleotide deletion, \=possible
CGDYSBERDÉPOVRERRÉCHIVFACESGACIP MSWTCDKEDDCEHGEDETHICHNEESGACIPE CONTRICISKOWLCDGSDDCGDGSDEAAHCE GKTCGFSSESCOTHEVYPERWL.CDGNDCA DCADESLAAGCLYSTCDDEFMCQNRQCIP HEFVODIBRDCADGSDESPECEYPTGGSESE RCANGRCLSSROWECDGENOCHDOSDEAPE NPHCTSPEHKCNASSOFLCSSGRCVAEALLCN GQDDCGDSSDERGCHINECLSRKLSGCSODC EDLIKIOFECRCRGFIRLEDGRTCADVDECS TTPPCSGRCNTHGSTSKLCVEGVATHACHOUNG HESCANTDEEPTLEARRYTLAKINLDGSNY TLKQGLANNAVALDFDYREQMITYDVTTO GSMIRRMHAGSNYQUHRTGLSNPDGLAV DWVGGNLYWCDKGGDTTEVSKLINGAYRTYL VSSGLREPRALVOVDQNGGTLEVSKLNGAYRTYL VSSGLREPRALVOVDQNGGTLEVSKLNGAYRTYL VSSGLREPRALVOVDQNGTLEVSKLNGAYRTYL KTLLISTILIREPROLIFYTLARCPDVPNHIPCK VNNGGSSCSNLCLLSFRGGRICKACPTNFYLGSD GRTCVSNCTASGPVCKNNCCIPFWWKCDTE DDCGGNSDEPPDCERFCCRGGQCCGTGICTN PAFICOGNNGCOGNGDEBREDCE VTCAPNGFCSIKRCTPRVWCDRDNDCVD GSDEPANGTCGNNCGGGDEBROCE VTCAPNGFCSGNCGGNSCERGRERCE VSTCAPNGFCCSGNCCHARWCCDGDHDCADGS DESCETFRCDMDGPCGNCGGGGEBROCE KNNCVPGRWQCDYDNDCGNNDSEPSCTPR PCSSESFSCANGRCIAGRWKCDGDHDCADGS DENCTFRCDMDGPCKSGHCHLRWRCDA DADCMDGSDEFACGTGYNTCDDEFQNNT LCRPLAWRCDGGDDCGNSDEEPECARPT QPTRCTNNBVCGNNSCCTRANGCC SSSLRCNMFDDCGDSSDEPECARPT QPTRCTNNBVCGNNSCCTRANGCC SSSLRCNMFDDCGDSSDEPECARPT QPTRCTNNBVCGNNSCCTRANGCC ARNTMKTHTCKACRGCRIATYCGNGCARPT NASCGGRGGGTGANSCC QPRYTGDKGCLUQCHYNGGGGCTLANSFCG MCTCRCTCONARDSCCTANSFCG MCTCRCTCONARDSCCTANSFCG MCTCRCTCONARDSCCTANSFCG MCTCRCTCONARDSCCTANSFCG MCTCRCTCONARDSCCTANSFCG MCTCRCTCONA	İ				sequence		nucleotide insertion
MSWTCDKEDOCEHGEDETHICNIKFCSEAQUE GONNICISKOWLOGSDECGESDEAAHCE GKTCOPSSPSCPGTHVCVPERWLCDGNECQL GKTCOPSSPSCPGTHVCVPERWLCDGNECQL BCADEALAGCL VINTSTCDREEMEQONROCCIP KHEVCDHENDCALGSBSESECEVPTCGPSEIF RCANGRCLSSROWCDGENDCHDSDEAPE NPHCTSPERKCNASSQFLCSSGRCVAEALL GONDOCGDSSDERGCHHECLSKISKISGCSODC EDLKIGFKCRCRPGFRLKDDGRTCADVDGSC TTFPCSGRCINTHGSVKCLCVEGYAPPGGDP HSCKAVTDEFPFLIFANRYVLRKLNLDGSNY TLLKCGLNNAVALDFDYREQMHYWTDVTTY GSMURRMHILNGSNVQVIHRTGLSNPDGLAV DWOGGNLJWVGDGRDTHEVSKLNGAYRTVJ VSSGLREPRALVVDVONCYLYWTDWGDHSIS IGRIGMDGSSSSYLYDTTLTIPWRGTLTLDVYTE RIYWADAREDYIEFASLDGSNRHVVI.SQDIPH IFALTI.FEDYYVWTDWTTTYTWRCH.TLD VYTE RIYWADAREDYIEFASLDGSNRHVVI.SQDIPH IFALTI.FEDYYVWTDWTTTYTWRYNCK. NNGOCSNLCLLSROGGHKCACPTWKCDTF DCGDHSDEPPDCPERFCRPQCPCCSTGCTO. RAPPORT OF THE CONTROL OF THE							RLGVKGVLFQPCERTSLCYAPSWVCDGAND
CONNECISKOWLCDGSDDEADED GRIDESDEADED GRADESIAAGCLYNSTCDREFMCQNROCIP KHEVCDHERDCADSDSERSECE-PYTCGPSEE RCANGRCLSSRQWECDGENDCHDQSDEAF NPHCTSPEHKCHASSOFICSGSRCVAEALLCN GODDCGDSSDERCCHENECLSKLSGCSQDC EDLKKIGFKCREPGFILKDDGRTCADVDBCS TIFFCSQRCHTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPFLIFANRYTLRKLINLDGSNY TLLKQGLNAVALDEDYREQWIN WTDVTTQ GSMRRMHILMGSNVQVLHRTGLSNPDGLAV DWGGNLTWCKRGRTIEVSKLNGAYRTVL VSSGLREPRALVDVQNGVTLYWTDWGTDST RIYWADAREDYIEFALSDNAFHAVILLDYTE RIYWADAREDYIEFASLDGSNEHVVLSQDIPH FALLTLEDYYVTDWETKSINRAHKITGITN KTLLISTLIRRMDLHVFILALROPDVPHHFCK VNNGGCSNLCLLSPGGGHKCACTHNYLGSD GRTCVSNCTASGFVCKNDKCPFWWKCDTE DDCGGHSDEPPDCEFFKCRPGGFCGTGICTN PAFICDGDNDCQDNSDEANCDHVCLSCGFK CTNTNRCIPGIFRCNQONCCDGEDERDCPE VTCAPNGPCSITKRCPRVWCDRDNDCVD GSDEFANCTOMTCGVDFERCKSGRCPFARW KCDGEDDCGGDSDEFKEECDERTCEPYOFRA KCDGEDDCGDSDEFKEECDERTCEPYOFRA KCDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGGCACTHAV CDGEDCGGGGGGGCACTHAV CDGEDCGGGGGGGCGGCGGCGGCGGGGGGGGGGGGGGGG							CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
GKTCGPSSTSCPGTHVCVPERWLCDGNDCQCIP KHEVCDHDRDCADGSDESPECEVPTCGPSEE KHEVCDHDRDCADGSDESPECEVPTCGPSEE RCANGRCLSSRQWCDGENDCHDQSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN GQDDCGDSSDERGCHNECLSRKLSGCGCVAEALLCN GQDDCGDSSDERGCHNECLSRKLSGCTAPALLCN GQDDCGDSSDERGCHNECLSRKLSGCTAPALLCN GQDDCGDSSDERGCHNECLSRKLSGCTAPALLCN GQDDCGDSSDERGCHNECLSRKLSGCTAPALCN GDDCGDSSDERGCHNECLSRKLSGCTAPALCN GDDCGDSSDERGCHNECLSRKLSGCTAPALCN GDDCGDSSERGCHNECLSRKLSGCTAPALCN GDDCGDSSTRUCHNEGDSPCTAPALCN TITLERGCLNAVALDEDYREQMITYTDVTIQ GSMIRRMHLMGSNVQULRIFGLSNPOLAV DWGGRLYWCDKGRDTIEVSKLNGAYRTYL SSGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTU VSSGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTU VSSGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTU TORGADAGSGRSSSVIVDTKITWPNGLTLDYVTU TORGADAGSGRSSRVIVDTKITWPNGLTLDYVTU VSGLREPRALVVDQNGYLYWDGWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTU TORGADAGSGRSSRVIVDTKITWPNGLTLDYVTU VSGLREPRALVVDQNGYLYWDGWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTU TORGADAGSGRSSRVIVDTKITWPNGLTLDYVTU TORGADAGSGRSSRVIVDTKITWPNGLTLDYVTU VSGLREPRALVVDQNGYLYMDGLTLDYVDGHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVDGHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVDGHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVDGHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVDGHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVDGHSL VNNGGCSNLCLSPGGGRGCTGCTTNSYLGSGL GTTCCCGDDCGCDGNDSDEANCHAPTROWACCO TORGADAGSGRCCATARVACCAGGGRGCCGGGCGGGGTAPTROWACCO TORGADAGSGRCCAGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGG							MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
DGADESIAAGCLYNSTCDDREFMCQNRCGIL HHPVCDHRDRCDAGDSESPECEPYTCGPSEIF RCANGRCLSSRQWECDGENDCHDQSDEAR NPHCTSPEHKCNASOGLCSSGNCVAEALLCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCPGFRLKDDGRTCADVDECS TTFPCSQRCINTIGSTKCLCVGGYAPRGGDP HSCKAVTDEEPFLIFANRYTLYKLINLDGSNY TLLKQGLNNAVALDFDYREQMIPVTDVTTQ GSMRRMHLMGSNVQVLHRTGLSNPDGLAV DWYGGNLYWCDKGRDTIEVSKLNGAYRTVVL VSSGLREPRALVYDVQNGYLYWTDWGDHSIL IGRIGMGSSSSSYDVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNPHVVLSQDIPH IFALTI-FEDYYVTDWTTOWTEKTSIRRAKITTGTIN KTLLISTLIBRMDLHVFIALRQPDVPNHFCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASOPVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRRGGPGTSTIGTIN PAFICDGDMSCGPPDCPEFKCRRGGPGTSTIGTIN PAFICDGDMSCDPDCPEFKCRRGGPGTSTIGTIN PAFICDGMDCQDNSDEANCDHVCLPSQFK CTNTNRCIGGRKCNQQDNCGDGEDERDCPE VTCAPNGFQCSTKRCIPRVWCDRDMCVD GSDEFANCTOMTGGVBERKCEDGERTCEPYQFRC KNNRCVGRWQCDYDMDCDNSDEESCTFR PESESFSCANGRCIABWNCDGDDDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGWKCDGDDDCADGS DEKDCTFRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGWKTCDEPGCNNT LCKPLAWKCDGEDDCGDNSDENFECCAPFY CPNRPFRCKNDRVCLWIGRQCDTDCOD GTDEEDCEPPTAHTTHCKDKEFLCROQCL SSLRCNMFDDCGORSDEESCIPEPGCNNT LCKPLAWKCDGEDDCGDNSDERFECCAPFY CPNRPFRCKNDRVCLWIGRQCDTBCQ GTDEEDCEPTAHTTHCKDKEFLCROQCL SSLRCNMFDDCGORSDEEDSCIPEPGCNNT LCKPLAWKCDGEDDCGDNSDERFECCAPFY CPNRPFRCKNDRVCLWIGRQCDTSCAGGHLYCQ QPCCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPBSAYEQAFQOBESVRDAMDVHYKA RQIDRGVTHLNISGLKMPRGIADWVAGNY WTDSGRDVEVAQMGGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLYQDNIQWPTCLAVDYNHSTYRA RQIDRGVTHLNISGLKMPRGIADWVAGNY WTDSGRDVEVAQMGGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLYQDNIQWPTCLAVDYNHSTICHAMDGT LRETLYQDNIQWPTCLAVDYNHSTICHAMDGT LRETLYQDNIQWPTCLAVDYNHSTICHAMDGT LRETLYQDNIQWPTCLAVDYNHSTICHAMDGT LRETLYQDNIQWPTCLAVDYNHSTICHAMDGT LRETLYQDNIQWPTCLAVDYNGERSCEVOK CSCLEGACVVNKQSGDVTCNCTDGRKACSHV CSCLEGACVVNKQSGDVTCNCTDGRKACSHV CSCLEGACVVNKQSGDVTCNCTDGRKACSHV CSCLEGACVVNKQSGDVTCNCTDGRKCEWN CSCLEGACVVNKQSGDVTCNCTDGRKCEWN CSCLEGACVVNKQSGDVTCNCTDGRKCEWN CSCLEGACVVNKQSGDVTCNCTDGRKCEWN CSCLEGACVVNKQSGDVTCNCTDGRKCEWN CSCLEG	ļ						CQNHRCISKQWLCDGSDDCGDGSDEAAHCE
KHFVCDHDRDCADGSDESPECEYPTICGFSER RCANGRCLSSRQWCDGENDCHDSDEAPK NPHCTSPERKCNASSQHLCSSGRCVAEALL, GODDGDSSDERGCHNECLSKLSGGSQDC EDLKIGFKCRCRPGFRLKDDGRTCADVDGG: TTFCCSQRCDHTBCSVCLVEGYAPBGGDP HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY TLLKQGLNAVALDEDYREQMITWTDVTTQ GSMRRMHLMGSNAVQULRTGLSNPDGLAV DWYGGNLYWCDKGRDTIEVSKLNGAYRTYL VSGGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYYTTE RIYWADAREDYIERASLDGSNRHVVLSQDIPH IFALTLFEDYYYWTDWETKSINRAHKTTGTN KTLLISTLHRPMLHVHLAKRDPYNHPCK VNNGGCSNLCLLSPGGHKCACPTNFYLGSI GRICGNDGSSRSVIVDTKITWPNGLTLDYYTT DDCGDHSDEPPDCPEFKCRFGQPQCSTIGCTN TALTHERDYYYWTDWETKSINRAHKTTGTN KTLLISTLHRPMLHVHLAKRDPYNHPCK VNNGGCSNLCLLSPGGHKCACPTNFYLGSI GRICVSNCTASQFVCKNDKCIPFWWCDTE DDCGDHSDEPPDCPEFKCRFGQPQCSTIGCTN AFICIGDDMCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNQODNCGDGEDERDCPE VTCARNQPQCSITRKGTPWWCDTDNDCDN GSDEPANCTOMTOGVDEFRCKDSGRCPARW KCDGEDDCGDOSDEPRECCDERTCPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSSSESPCAMBGRCAGRWKCDGDDCADG GSDEPANCTOMTOGVDEFRCKDSGRCPARW KCDGEDDCGDOSDEPRECCDERTCPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSSSESPCAMBGRCAGRWKCDGDDCADG DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDEBSCTPR PCSSSESPCAMBGRCAGRWKCDGDDFECADG DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDEBSCTPR PCSSSESPCAMBGRCAGRWKCCDGBPECADG DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNRV CPPRRFPRKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTHCKNKKEFLCRNQRCL SSSLRCNMFDDCGDGSSDEBDCSIDPKLTSCAT NASICGDEACVTRKAAYCACRSGFVIVYADDNERS LFPGHPISAYEQAFQGDESVRDAMDVHYKA GRVYWNWHTGTTHTHCKNKKEFLICKMQRCH RCIQDAMADGSSUSTRDAMMPRETAMADG LKETLVQDNIQWTGLAVDTHNERLYWADA LKLSVIGSRUNGTDPVAADSKRGLSHFFSIDO FEDVIYOTYNNWFKHKRGRSLYPFSTDP CLSPGPWCTCPMGRKLDNGTCANSPSDP CLSPGPQCTCPMGRKLDNGTCACRSCCVX GRYTGDKCELDQCWEHCRNGGTCAASPSQ MMTCRCTGTGTWCTLQVCAQCACNTSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCT VNQGNQPQCRCLPGRLDDCQVACANNSTCM CSRCLBGACVVNKQSGDVTCNCTDGRCAPN	1						GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
RCANGRUSSRQWECDGENICHIOSIDEARY NPHICTSPEHKCNASOFICESSRCVAEALLCY GODDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRGFRLKDDGRTCADVDECS TITPFCSQRCNTHOSYKLLCYGYAPRGGDP HSCKAVTDEBFFLIFANRYTLRILNLDGSNY TLLKQGLNNAVALDFDYREGMIYWTDVTTQ GSMIRRMHLNGSNYQVLHRTGLSNPEDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMGGSSRSVIVDTKITWFNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYTDWETKISNRAHKTTGTN KTLLISTLIRPMDLHYFIIALRQPDYPNHPCK VNNGGCSNLCLISPGGGHKCFTFYVLGSD GRTCVSNCTASQFVCKNDKCFFFWWKCDTE DDCGGMSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLFSQFK CTNTNCHGFIFRCNGQDNCGDGEDEDCPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDEDECPE VTCAPNQPCGSTIKRCHPRWCDRDNDCVD GSDEPANTOMTOCYDEFRCNOGONGCDETPOCHT CKNNRCVPGRWQCDYDNDCGDNDDCCFTARW CCDEDDCGDGSDEDECFTCPTQFNT LCCPLAWKCDGEDDCGDNSDENTECARPV CPPNPFRCKNDRVCLWIGGCDGTDNCGD DECCTPRCDMDQPCCKSGFICPILRWRCDA DADCMGSDEEACGTGVRTCDGPCNNTT LCCPLAWKCDGEDDCGDNSDENTECARPV CPPNPFRCKNDRVCLWIGGCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFILCRNQRCL SSSLRCNMFDDCGGSDEDCHTCAGGL SSSLRCNMFDDCGGDSDEDDFLTSCARPV CPPNPFRCKNDRVCLWIGGCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFILCRNQRCL SSSLRCNMFDDCGGSDEDCHTTYTON RANSICGDEARCVATERAAVCACRSGFITYPG QPCCQDINECLRGTGTSQLCNMTXGGHLCSS ARNYMKTHNTCKAEGSEYQVLYIADDNERS LPPGHPHBATEQAFQGDGSVTLNSGGLTVOG QPCYWTNWTIGTTSTRSLPAPTTSNRRR RQDRGTVHLNISGLKMRGGADWVAGNYY WTDSGRDTUPCAQNKGCRGGSCFLNARRQPKCCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC QPRYTGNKCTLNNRVFKHIRFGTLAASPCKCC RFGTCQMAADGSCRCTAYFGGSCFLNARRYPCCC RFGTCQMAADGSCCTANSKMMPCCQCPHM TCPRCEENTSQQQPGHASILIPLLLLLLL							DGADESIAAGCLYNSTCDDREFMCQNRQCIP
NPHCTSPEHKCNASSGRCVAEALLCN GQDDCGBSSBERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRFGFRLKDDGRTCADVDECS TTFPCSQRCNTHGSYKCLCVEGYAPRGDP HSCKAVTDEBFFLIFANRYYLRINLDGSNY TLKQGLNNAVALDFDVREQMIYWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYKTVL VSSGREPRALVDVQNGYGLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIFFASLDGSNYHVLSQDIPH IFALTLFEDVYWTDWETKSINRAHKTTGTN KTLLISTLIERPMDLHYFILALRQDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTFYLJGSD GRTCVSNCTASGPVCKNDKCIFFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQPGCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLFSGFK CTNTNRCHGGIRGCNGQDNCGDGEDERDCPE VTCAPNQPQCSITKRCIPRVWVCDRDNDCVD GSDEPANTGVMTCOVDEFRCKDSGRCIPARW KCDGEDDCGDSSBEPKECDERTCEPYOFRC KNNRCVPGRWCQDYDNDCGDSBEDCFE VTCAPNQPQCSITKRCIPRWWCDRDNDCVD GSDEPANTGVMTCOVDEFRCKDSGRCIPARW KCDGEDDCGDSSDEPKEECDERTCEPYOFRC KNNRCVPGRWCQDYDNDCGDNSDEESCTPR PCSSEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCXSGRCIPARWCDA DADCMDGSDEAACGTGVRTCPLDEPQCNNT LCKPLAWKCDGEDDCGDNSDEPSCTPR CPPNEPRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFILGNGRCL SSSLRCMMFDDCGDGSDEDFKLTSCAT NASICGDEARCVRTEKAAVCACRSGFHTIVPG QPCQDNIECLRGTGTSGLCNTKGGGHLCSC ARPHMKTHNTCKAGSEFYQNLYMADDVIRKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRIR RQDDRGAYHLINGGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGRDVETAAMSGLKMPRGAIDDWAGNYY WTDSGROVTENNRYKEHKFENLYNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSSSPCYCTOPYNGKTLDNGTCVPYPSTPP PDAPPGTGCLAQCYNGGGCFLNARROPKCC QPRYTGMCCELOPGGSGCFLNARROPKCC QPRYTGMCCTIOPHCRATGGCAASPSG MPTCCCTTOPHCSTNGGCTCAASPSG MPTCCCTTOPHCSNGGGCTLNARROPKCC VNGGNOPQCGCGCGCE NFGTCQMAADGSRQCCCTAYFEGSRCEVNK CSRCLEGACVVNKCGGGOTTCNCTDGCAYPS CLTCUGHCSNGGGTAMSKMMPECQCPPHM							KHFVCDHDRDCADGSDESPECEYPTCGPSEF
GODDCGDSSDERGCHINECLSRKLLSGCSQDC EDLKIGHKCRCRGPGHLADDCGS TITPCSQRCINTHGSYKCLCVEGY APRGGDP HSCKAVTDEEPFLIFANRYYLRLINLIDGSNY TLLKQGLNNAVALDFDYREQMYWTDVTTQ GSMIRRMHLNGSNYQVLHRTGLSNPEDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMGGSSSKJVDTKITWPGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLEEDYVYTDWETSINRAHKTTGTN KTLLISTLIRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCYSNCTASGYVCKNNSCHPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFCCSTGICTN PAGFICDGDNDCQDNSDEARGOPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFCCSTGICTN PAGFICDGDNDCQDNSDEARGOPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFCCSTGICTN PAGFICDGDNDCQDNSDEARGTPWWCDRDDCVD GSDEFANCTQNTCGVDEFRCKDSGRCPARW KCDGEDDCGDGSDEPKECRPGGFCCSTGICTN PAGFICDGDNDCGDSDEFKERCPYOFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSSEFSCANGRCHAGRWCCDGDHDCADGS DEKDCTPRCDMDQFQCKSGRCPARW KCDGEDDCGDBSDEPRECCARV CPPNEPFRCKNDRYCLWIGRQCDGTDNCGD GTDEBCCEPFTAHTTHCDKKFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCYNTEKAPAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYLADDBERS LFPGHPHSAYEQAFQGDGSDENRIBR RQDDRSCHTEKAPAVCARSGFHTVPG QPGCQDNECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYLADDBERS LFPGHPHSAYEQAFQGDGSVKIDAMDVHYKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRHR RQDDRSVTHINSGLKMPRGIADWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGHBYRETAAMDGT LEETLVQDNQWPTGLAVDYHNERLYWADA KLSVIGSBLNGTDPIVAADSKGLSIPPSDV FEDYIYGVTYNNRYFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQFEVTNPCDKKCGWL CLLSSSPCYCTPONKRLDINGTCVPYPSPTPP PDARPGTCNLQCFNGGSCFLNARRQPKCCC QPRYTGDKCFLOQCCAGYCANNSTCT VNQGNQPQCRCLPFGTGRGCCAASPSG MPTCRCPTGFTGFKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPFGTGRCQCCGSYCC NFGTCQMAADGSRQCRCTAYFGGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMASRMMPECQCPPHM TGPRCEEHVFSQQQFGHASILIPLLLLLLLLLLLLLV VAGVVFWYKRRYQGAKGFGRTMSGRMMPECQCPPHM TGPRCEEHVFSQQQFGHASILIPLLLLLLLLLLLV VAGVVFWYKRRYQGAKGFGRTMSGRMMPECQCPPHM TGPRCEEHVFSQQQFGHASILIPLLLLLLLLLLLLV VAGVVFWYKRRYQGAKGFGRTMSGRMMPECQCPPHM							
EDI.KIGFKCRCRPGFRILKDDGRTCADVDECS TTFPCSQRCINTINGSYKCLVOPAPRGGDP HSCK.AVTDEEFFIJFANRYYLRXLNI.DGSNY TLIKQGLNAVALDFDYREQMIYWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNI.YWCDKGRDTIEVSKLNGAYRTVL VSSGR.REPRALVOVQNGYLY-WTDWGDHSI. IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNYLVLSQDIPH IFALTI-EEDVY-WTDWETKSINRAHKTTGTN KTLLISTLIRPMDLHYFILALGPDPYNHPCK VNNGGCSNLCLLSFGGGHCGACPTFFYLGSD GRTCVSNCTASGFVCXNNEGPWWKCDTE DDCGDHSDEPPDCPEFKCRPGQPCGSTGICTN PAFICOGIONDCQDNSDEANCDIHVCLPSQFK CTNTNRCHOFFRCNGQNCGDGEDERDCPE VTCAPNQPQCSTTRKCRPRWVCDRDDDCVD GSDEPANTGOMTCOVDEFFCKDSGRCPFARW KCDGEDDCGDSSDEPKECCDERTCEPYOFFC KNNRCVPGRWCQDYNDNCGDNSDEESCTPR KCDGEDCGDSSDEPKECCDERTCEPYOFFC KNNRCVPGRWCQDYNDNCGDNSDEESCTPR PCSSEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPFCCDMDQPCXSGRCPLARW KCDGEDDCGDDSDEPNECCARPV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCGPTATTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDIPFLTSCAT LCKPLAWKCDGEDDGGNSDEFECCARPV CPPNRPFFCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGGSSEEDIPFLTSCAT NASICGBEARCVRTEKAAYCACRSGFHTIVPG QPCQDNINGLRFGTGSSGLONTKGGHLCSC ARPHMKTHNTCKAEGSEYQVLYLADDNEIRS LPPGHPHSAYEQAFQODESVRDAMDVHVKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRIR RQIDDGAVTHNWHTGTISYRSLPPAAPPTTSNRIR RQIDDGAVTHNWHTGTISYRSLPPAAPPTSNRIR RQIDDGAVTHNWHTGTISYRSLPPAAPPTSNRIR RQIDDGAVTHNWHTGTISYRSLPPAAPPTSNRIR RQIDDGAVTHNWHTGTISYRSLPPAAPPTSNRIR RQIDDGAVTHNWFTGTMYWSDWGNIPKIETAAMDGT LREILVQDNIQWFTGLAVYHNDRGTCVPYPSTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKCRCDNTCTVPVPSTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKCRCDNTCTVPVPSTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKCRCDNTCTVPVPSTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKCRCDNTCTVPVPSTTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKCRCDNTCTVPVPSTTPP PDARPGTGNLQCFNGGSCFLNARRQPKCCC QPRYTGNCCHOPKGGGCFLNARRQPKCCC QPRYTGNCCHOPKGGGTGCASPSG MPTCCCTTOPKCTQQVCAQYCANNSTCT VNQGNQPQCCCLPFIGMGCTCAASPS CLTCCGGCGCFLNARRQPCCCCFMM CSRCLGGACVVNKQGGNTCANTYSG	1						NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
TIFPCSQRCINTHGSYRCLCVEGYAPRGGDP HSCKAVTDEEPFLIFANRYLRIALDGSNY TLLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNFDGLAV DWVGGNLYWCDKGRDTEVSKLNGAYKTVL VSSGLREPRALVYDVQNGYLYWTDWGDHSL IGRIGNGMGSSSSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALT-FEDYYWTDWETSNRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHACOPHYNHCK VNNGGCSNLCLLSPGGGHACOPHYNHCK VNNGGCSNLCLSPGGGHACOPHYNHCK VNNGGCSNLCLSPGGGHACOPHYCLPSQFK CTNTNRCIPGIFRCNQQDNCGDGEDERDCPE VTCAPNQFCQSTIRKCIPRVWCDRDNDCVD GSDEPANCTQMTCGVEFRCKDSGGCPARW KCDGEDDCGDSDEPKECBRTCGPYQFEC KNNRCVPGRVWCDYDNDCGDNDEESCTPR PCSSEFSCANGRCIAGRVCWCDGDHDCADGS DEKDCTPRCDMQCFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCODNSDENFEECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPTPATHTHCKREFICRNQRCL SSSLRCNMFDDCGGDSDEPKECBCNGRCL SSSLRCNMFDDCGGDSDEDCSIDPKLTSCAT NASIGGDBEACVYTEKAACCRSGHTTVP QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLVIAGDNIERS LFPGHHSAYEQACGOGSDRIANDVHVKA GRYYWTNWHTGTISYRSLPPAPPTISNRHR RQIDRGVTHLNISGLLMRPGIADDWAGNVY WTDSGRDVEVAQMKGENKTLISGMDEPH AIVVDPLKGTMYWSDWGNIPKIETAAMDGT LRETLYQDNIOWPTGLAVDYINERLYWADA KLSVIGSRLAMINGTOLPVPSPTPP PDAPRRGTCNLQCTRGGSCTLNARGPKCCW CREATIVE STANDAY CLESSEGVCTOLNGCGCCHANNETCG ULSSEGVCTONGKGCLDGCCCANNSTCG QPRYTGDKCGLDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQOVCAGYCANNSTCT VNQGNOPGCLLPGFLGDRCQNRCCSWCC VPRYTGDKCLDQCCCCANNSTCT VNQGNOPGCLLPGFLGDRCQNRCCSWCC VPRYTGDKCCLDQCCCANNSTCT VNQGNOPGCLCPGTLORGCCCANNSTCT VNQGNOPGCCLPGFLGDRCQNRCCSGCCC VPRYTGDKCCLDQCCCANNSTCT VNQGNOPGCCLCPGTLORGCCCANNSTCT VNQGNOPGCCLCPGTLORCCTORCVAPS CLTCVGHCSNGGSCTLNARKMPECCCPPHM TGPRCEEHYFSQQQPGHIASILLILLLLLLL VAGVVPVKKRVQQAKGGGPQCRATNTEGAM CCLTCVGHCSNGGGCTLNARKMMPECCCPPHM TGPRCEEHYFSQQQPGHIASILLILLLLLLLV VAGVVFVKRRVQQAKGGPQCRMTTGAM	1						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
HSCKAVTDEEPFLIFARRYYLRIKLINGSNY TILKQGLNAVALDEPYREOMIYWTDVTTQ GSMIRRMHLNGSNYQVLHRTGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVPDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEPASLDGSNRHVVLSQDIPH IFALTLEPDYYWTDWETKSINRAHKTGTN KTLLISTLHRPMLHYPHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCPFWKCDTE DDCGDHSDBEPPDCEPEKRCPGQCSTGICTN PARICOGNDCQDNSDEANCDHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCTPRVWVCDRDDCVD GSDEPANCTOMTGVDEPFCCNSGRCIPARW KCDGEDDCGDDSDEPKEECDERTCEPYQFRC KNNRCYPGRWQCOYDNDCGDNSDESCTTPR PCSESERSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPILWRCDA DADCMDGSDEEACGTGVTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRFFRCKDRVCVGWGCOTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASIGGDEARCVTREKAYACACRSGFHTYPG QPCCQDINECLRFGTGSQLCNNTKGGHLCSC ARNFMKTHNTCKAGESEYQULVIADDNEIRS LFFGHFBSAYEQAFQODESVRIDAMDWHVKA GRYYWTNWHTGTISY SIZAAPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRKLISGMDEPH AIVYDPLAGRTMYWSDWGNIPKIETAAMDCT LRETLYQDNIQWPTGLAVDYNERILYWADA KLSYIGSRLANTWYFGTSYSRJAAPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRKLISGMDEPH AIVYDPLAGRTMYWSDWGNIPKIETAAMDCT LRETLYQDNIQWPTGLAVDYNERILYWADA KLSYIGSRLANGTDPIVADARGLSSHFSIDV FEDYIYGVTYNNRVFKHKKFGHSPLVNLTGG LSHASDVVLYHOHKQPETNCDRKXCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRRGTCNLQCFRGGSCFLNARRQPKCC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPRCTQQVCAGYCAGNSTCT VNQGNQPGCLLPG-TGORCYNRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVWNKQGGDVTCNTCDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHYFSQQQPGHASILIFLLLLLLLVL VAGVVFWYKRRVQCAAKGFQIQRMTNGGAM	1						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
TLLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMRRMHLMGSNVQUHGLSWFDGLAV DWVGGNLWCDKQRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLLYWTDWGDHSL IGRIGMGSSSSNVIDTKITYPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRIHVYSQDIPH IFALTLFEDYYWTDWETKSNRANETTGTN KTLLISTLHRPMDLHVFHALRQPDYNHPCK VNNGGCSILCLLSPGGGRIKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPQFQCSTGICTN PAPICDGDNDCQDNSDBANCDIHVCLPSQFK CTNTINKCIPGIFRCNOQDNCGDGEDEBCCPE VTCAPNQFCQSTIKTCIPRYWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCPARW KCDGEDDCGDSDEPKEGDRTCEPYQFRC KNNRCVPGRWQCDVDNDCGDNSDESCTPR PCSESFSCANGRCIAGRWKCDGDHOCADGS DEKDCTPRCDMDOFQCKSGHCPLRWRCDA DADCMDGSDEBACGTGVTVTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPTPAHTTHCKACKEFLCRNQRCL SSSLRCNMFDDCGDOSDEEDCSIDPKLTSCAT NASIGGDBEACGTGSQLCNNTKGGHLCSC ARNPMKTHNTCKAGSSEYQULVJADDNERS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYYMTNWHTGTIS YRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPAAIDUVAGNVY WTDSGRDVEVAQMKGENRKLISGMDEPH AIVVDPLRGTMYWSDWGHIPKIETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNUFFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KINGTDYPVAGNIFKETAAMDGT LREIL VQDNIGWPTGLAVDYNRERLYWADA KLSVIGSRIL KORTONTORTORTORTORTORYBPTP PDAPRRGTCNLQCFNGGGCFLNARKCEWL CLLSSEGFCCHARCAGGGCTNARKMMPECCCPTGFTTPRCTQQCAGYCAGNTSCT VNQGNOPGCLLPGCLEDGCCGCTAATSTGCT WTCGCCPTGFTTTRCTCTDGCCPTH TGRCCEHYSGQCGCTAATSTCT VCGGCGCTAATSTCGCCPTHMTGGAAMTGAAN							TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
GSMIRMHLINGSNVQVLHRTGLSNPGLAV DWGORLYWCDGROTEVSKLINGAYRTVL VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMGSSRSVIVDTKITTWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLEPDYYWTDWETKSINRAHRTGTN KTLLISTLIRRMDLHVPHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTFYLGSD GRTCVSNCTASQFVCKNDCIPFWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PARICDGDNDCQDNSDEANCDIHVCLSPGFK CTNTNRCIPGIFRCNQGDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRDNDCVD GSDEPANCTOMTGVOYDETRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYOFRC KNNCVPGRWQCDYDNDCGDHSDENDECVE VTCAPNQFQCSITKRCIPRVWCDGNDDCVD GSDEPANCTOMTGVOYDETRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYOFRC KNNCVPGRWQCDYDNDCGDHSDEPECSTPR PCSSEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTTRCDMDGFQCKSGHCPLRWRCDA DADCMDGSDEAGGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGGDSGDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGHTTVPG QPGCQDINECLRFGTGSQLCNNTKGGHLCSC ARNFMKTHNTCKAGESEYQVLVIADDNEIRS LFFGHFBSAYEQAFQGDESVRIDAMDVHVKA GRYVMTNWTGTISYRSJCHAPAPFTTSNRJR RQIDRGVTHLINISGLKMPRGIALDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIEKIETAAMDGT LRETLVQDNIQWPTGLAVDYHIRELYWADA KLSYIGSRLNGTDPVAADNIERSTJWV ROUNDRYCTLANDTHERSTJWV REDYTGNITWYSDWGNIEKIETAAMDGT LLRETLVQDNIQWFTGLAVDYHIRELYWADA KLSYIGSRLNGTDPVAADRGLSHFSIDV FEDYIYGVTYNNRVKHIKFGHSPLVNLTGG LSHASDVVLYHQHKQPCTPORDRKCKEGMI CSRCLEGACVWNKQSGUNKTCHDCRXARSG MPTCRCPTGFTGPRCTQQVCAGYCANNSTCT VNQGNOPGCRCLPGFLGBRCCNNTTCRAPS CLTCYGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHYFSQQQPGHASILIFLLLLLLLVL VAGVVFWYKRVQQAKGFQRGCMTNTGAM	1						
DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRAL VVDVQNYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDVIEFASLDGSNRHTVLSQDIPH IFALTLFEDYYYWTDWETKSINRAHKTTGTN KTLLISTLIRPMDLHVFHALRQDPVPNHPCK VNNGGGSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGPGSTGICTN PAFICDGDDVCQDNSDEANCDHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPWVCDRDNDCVD GSDEPANCTQMTCGYDEFRCKBGRCPARW KCDGEDDCGDSSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDDSDEESCTFR PCSSSEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTFRCMDDPGCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECAFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCKNQRCL SSSLRCNMFDDCGDGSDEEDCSDFKLTSCAT NASICGDEARCVXTEKAAYCACRSGFHTVPG QPGCQDINECLREFTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDMADVHYKA GRYVWTNWHIGTISTYSLPPAAPPTTSNNHR RQIDRGVTHINISGLKMPRGIAIDWAGNVY WTDSGRDVIEVAQMKGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNEPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPPSIDV FEDYIYGVYTNNNFYKHKRGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNEPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPPSIDV FEDYIYGVYTNNNFYKHKRGENPLVNLTGG SLSHASDVULTHQHKQPGCFPTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCAPSPSOY MPTCRCTGTGTGFKCTQQVCAGYCANNSTCT VNQGNQPCCLLOFIGGGCFLNARROPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSQ MPTCRCTGTGTGFKCTQQVCAGYCANNSTCT VNQGNQPCCLLOFIGLGRCQYRCCSGYCE NFGTCOMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNNCGACGOTCNCTCTGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHABSILLPLLLLLLLULV VAGVVPWYKRRVQQAGFQRCTAYFEGSRCEVNK							
VSSGLREPRALVVDVQNGYLYWTDWGDHSL IGRIGMDGSSRSVIDTKITWPNGLTLDYVTE RIYWADAREDVIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTTGTN KTLLISTLIRPMDLHVPHIALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGPGCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDEEDCPE VTCAPNQFQCSITKRCIPRWVVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEFKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSSESFSCANGRCIAGKVCDGHDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEAACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDEEDCSIPFLTRCDADG GTDEEDCEPFTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIPFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPQ QPGCQDINECLREGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYADDNEIRS LFPGHPHSAYGAAFQDGSVRIDAMDVHVXA GRVYWTNWHTGTISVRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNIPRETAAMDGT LREITLVQDNQWFTGLAVDVHNERLYWADA KLSVIGSRLNGTDFVADSKRGLSHPPSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKGPCVPVPSPTPP PDAPRFGTCNLQCFNGSCFLNARRQFKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQVCAGYCANTSCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKGGGCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILLPLLLLLLLULL VAGVVFWYKRRVQQAGAKGFQHRNTTNGAM							
IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLOGSNRHVVLSQDIPH IFALTLFEDYVYWTDWTKSNRAHKTTGTN KTLLISTLIREMDLHVFHALRGPDVPNHPCK VNNGGCSNLCLLSRGHCACPTNFYUGSD GRTCVSNCTASGPVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPERKGRGGPCCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCCDGEDERDCPE VTCAPNOPCQSTIFKCIPFWVVCDRDNDCVD GSDEPANCTQNTCGYDEFRCKDGRCPARW KCDGEDDGGDSSDEPKEECDERTCEPY VPRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSSEFSCANGRCIAGRWKCDDGHDCADGS DEKDCTFRCDMDQPCCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCJDEFQCNNT LCKPLAWKCDGEDDGCDNSDENEECAFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVKTEKAAYCACRSGFHTVPG QPGCQDIDNECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEVQVLYJADDNERS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYVTNWHTGTISTYSLPPAAPPTTSNRHR RQIDRGVTHINISGLKMRGIAIDWAGNVY WTDSGRDVIEVAQMKGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNIPKETAAMDGT LREITLVQDNIQWFTGLKYTSLLPAADGT LREITLVQDNIQWFTGLAVYHPELLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRYFKHHKFGHSPLVNLTGG SLSHASDVULTYGHKEPGSCFLNARROPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTTGTTGFKCTQQVCAGYCANNSTCT VNQGNQPCCLLDGT,GDRCQYRQCSGYCE NFGTQOMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGVCTCNTDGSRVAPS CLTCVGHCSNGGSCTIMNSKMMPECQCPPHM TGFRCEEHVFSQQQPGHABSILLFLLLLLLLLULV VAGVVPYWRRRVQQAKGFQRRMTNGAM							
RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IIFALTLFEDYVYWTDWETKSINRAHKTTGTN KTLLISTLIRRPMULVFIHALRQPDVPNHPCK VNNGGCSNLCLLSPGGHRKACPTINFYLGSD GRTCVSNCTASGPVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRDNDCVD GSDEPANCTQMTCGDFFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDERTCEPPQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIGARWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGNCDGTDNCGD GTDEEDCEPPTAHTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDSIDPKLTSCAT NASICGBEARCVATEKAAVCACRSGHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQULYJADDNEIRS LFGHPHSAYEQAFAQDESVRIDAMDVHVKA GRYYMTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVHLNISGLMPRGIADDWAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHRKIETAAMDGT LRETLVQDNIQWPTGLAVDYHRKETLYWADA KLSVIGSRLINGTDPIVAADSKRGLSHPPSIDV FEDYIYGVTYNNRVFKLIKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTGTNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGLGBRCQVTRQCSGYCE NFGTCQMAADGSRQCRCTAYFGSRCEVNK CSCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTINNSKMMPECQCPPHM TGPRCEEHVFSQQQFGHIASILIPLLLLLLLLLL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ţ						`
IFALTLEDVYWTOWETKSINRAHKTTÖTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGHKCACTTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIERCNQQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRDNDCVD GSDEPANCTQMTGGVDEFRCKDSGACCIPARW KCDGEDDCGDGSDEPKECDEPTCPYQFFC KNNRCVPGRWQCDVDNDCGDNSDEESCTPR PCSSSEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTIDNCGD GTDEEDCEPPTAHTTHCKDKREFLCRNQRCL SSSLRCMMFDDCGDSDEEDCSIPFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTHANTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRTLISGMIDEPH AIVVDPJLRGTMY WORGHPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYNNRVEKHKFGGHSLYNLTGG LSHASDVVLYHQHKQFEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKGCTLAARQPKCRC QPRYTGDKCELDQCWEHCRGGTCAASPSG MPTCKCTPGTFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAAYFGSRCEVNK CSCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTINASKMMPECQCPPHM TGPRCEEHVFSQQQFGHIASILIPLLLLLLLLL VAGVVFWYKRRVQGAKGFQHRKMTNGAM	i						
KTLLISTLIRPMOLHVFHALROPDVPNIFICK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASGPYCKNDKCIPFWWKCDTE DDCGDHSDEPPDCEFKCRPGGFQCSTGICTN PAFICOGDNDCQDDBCANCDIHVCLPSOFK CTNTNCIPGIFRCNGQDNCGDGEDERDCPE VTCAFNQFQCSITKECTRRVWCDRDNDCVD GSDEPANCTQMTCGVDEFFCKDSGRCTPARW KCDGEDDCGDGSDEPRECDERTCEPYQFRC KNNRCVPGRWQCDYDDCGDNDESSCTPR ENDERSTER PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGGIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDSIDPKLTSCAT NASICGDEARCVATTEKAAYVACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSSEYQVLYIADDNEIRS LFPGHPHSAYEQAFGADSVADDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRXTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHRERLYWADA KLSVIGSIRLINGTDPIVAADSKRGLSHPPSIDV FEDYIYGVTYNNRVEHKFKGHSLYNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTGTNLQCFKHKFKGGSCTLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFGSRCEVNK CSCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTINNSKMMPECQCPPHM TGPRCEEHVFSQQQFGHIASILIPLLLLLLLLLL VAGVVFWYKRRVQAGKGFQHTQRMTNGAM			.			i	·
WNNGGCSNLCLLSPGGGHKCACPTIFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFECNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFTCKDSRGCPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDCGDNSDEPECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGGSSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYJADDNERIS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWRTSYRSLPPAAPPTTSNRH RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRAVETAVAQKMGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNIFRKETAAMDGT LRETLVQDNIQWPTGLAVDYHERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRRGTCNLQCFNGGSCTLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRGSCTLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCCPTGFTGRCTQVACASPSG MPTCCCPTGFTGRCTQVACAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSCCLEGACVYNKQSGDVTCNCTTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLV VAGVVFWYKRRYQGAKGFQHQRMTNGAM							
GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPQGPCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLFSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEEDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMGSDEEAGGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFFRCKNDRVCLWIGRQCDGTIDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVTEKAAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNSGLKMPRGIAIDWVAGNVY WTDSGRVEWAMGGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWFTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHFPSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQVRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVIK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLV VAGVVFWYKRRVQGAKGGFQRTNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLLLV							•
DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNDEANCDIHVCLPSQFK CTNTNRCIPGIFECNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRDNDCVD GSDEPANCTOMTCGVDEFRCKDSGRCPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPJLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFECCARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVETEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQCDBSVRIDAMDVHVKA GRVYWTNWHTGTISVRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIALDWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHRQPEVTNPCDRKKCEWL CLLSPSGPVCTPOKKLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGGKCTQVVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNRQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLV VAGVVFWYKRRVQGAKGFQRGTNSAMM							
PAFICOGONDCODNSDEANCDITYCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNOFQCSITKRCIPRYWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGYKTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNEPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSBEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYJADDNEIRS LFPGHPHSAYCAPQGDESYNDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLYQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVALTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCRLQCPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTINNSKMMPECQCPPHM TGPRCEEHVFSQQPGHIASILIPLLLLLLLLLLL VAGVVFWYKRVQGAKGFQHQRMTNGAM							3
CTNTNRCIPGIRRCNGODNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWCDRNDCVD GSDEPANCTQMTGQVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDVDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGWRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHKCDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPCQDINECLRFGTCSQLCNNTKGGHLCSC ARNPMKTHNTCKAEGSEFYQVLYJADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIADWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDUGNIEPKETAAMDGT LRETLVQDNIQWPTGLAVDVHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCYNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADSRQCRCTAYFGSSCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQOFGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							• • •
VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGGGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPHSAYEQAFQDESVRIDAMDVHVKA GRYYWTNWHTGTISVRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGBCCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPFGLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHYFSQQQPGHIASILIPLLLLLLLLVL VAGVVFWYKRRVQGAKGFQHRMTNGAM			•				
GSDEPANCTOMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGWTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKNKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECREFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPFSDV FEDYIGVTYINNRVFKHKKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
KCDGEDDCGDGSDEPKEECDERTCEPYOFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSSESFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGBEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEVQVLYJADDNEIRS LFPGHPASAYCAGPGOESVRDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIEPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPTGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTIMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							, ,
KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSCHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIFPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTIGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGSCTIMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							l
PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPILRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVWEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKOPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTTPP PDAPRFGTCNLQCFNGGSCFLNARROPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLITCVGHCSNGSCTMNSKMMPECQCPPHM TGPRCCEHVFSQQQPGHIASILPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQODESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			1				
DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDGGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSUGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFKCTQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGSCTIMNSKMMPECQCPPHM TGPRCEEHYFSQQOPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1		- 1				
LCKPLAWKCDGEDDCGDNSDENPEEČARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNFVKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHYPSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVPKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			· ·				· 1
GTDEEDCEPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARRPGTCNLQCFNGGSCFLNARQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQOPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			i		Ì		7
NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			- 1	ł			
QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCKTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						
ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1 '		Į.				
LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			ľ]		
GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM						İ	
RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			Į		į		
WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				1			
AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			- 1				
FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						LRETLVQDNIQWPTGLAVDYHNERLYWADA
LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1	1	- 1	1	ļ	l	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	j	1				l	FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG
PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM		1		į			
QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM						1	
MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM						l	
VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1 1	- 1	ŀ	ł		į	
NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			I	ļ			
CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			ł	1			
CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						
TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	}						CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1 1		ŀ	l			
						ŀ	
NVEIGNPTYKMYEGGEPDDVGGLLDADFAL	1 .		}				
	L		l				NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

NO. of No. of No. of No. of n	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence USSN Docation Docation Sequence Sequence Docation Sequence Do	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence 93496 corresponding on the first of the side decidence peptide scale residue of peptide sequence peptide seque		1		_	•		
union coling to first amino acid residue of peptide sequence pepti			i				
mino acid residue of peptide sequence peptide s		uence		l .			
	uence			714			
Populida Populida				i			
	j		1			soquense	
DPDKPTNFTNPVATLYMGGHGSRHSLASTD			ŀ			Ī	
			 -		1		
AGILPIPTIPNI-TILGYSISSIGAIPAAALDPI ATLGEPQPI-MONDPSKIDEIRTY-YVONI ATLGEPQPI-MONDPSKIDEIRTY-YVONI NSQITTADOLLEFFKQVGEVKFVRMAGDET QPTRTA-FVER-ADONS-VPALA-HONG-WEGDPP LKINISINNAIVKPPEMTPQAA-AKELEEVMKR VREAQSFISAAIEPGWH-STSTLOGF*-IRR RWTRP-PACTICGTFHILCLIINWDL-PLP-ATTA K*FPPPKVWKEQ*KKRRKSRSHTRSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSK			1				
AGILPIPTIPNI-TILGYSISSIGAIPAAALDPI ATLGEPQPI-MONDPSKIDEIRTY-YVONI ATLGEPQPI-MONDPSKIDEIRTY-YVONI NSQITTADOLLEFFKQVGEVKFVRMAGDET QPTRTA-FVER-ADONS-VPALA-HONG-WEGDPP LKINISINNAIVKPPEMTPQAA-AKELEEVMKR VREAQSFISAAIEPGWH-STSTLOGF*-IRR RWTRP-PACTICGTFHILCLIINWDL-PLP-ATTA K*FPPPKVWKEQ*KKRRKSRSHTRSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSKSK	374	1724	Α	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG
NSGITTADQLLEFFKQVGEVKFVRMAGDET	ł		ļ		1		
	1					1	ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL
LKINHSNNAIVKPPEMTPQAAAKELEEVMKR VREAQSFISAAEPGVLHSTOPHIGGPTSR RMYRE*APCTICGTFHICLINWDD-I_P-AYTA K*FFPRVWKEQ*KKRRKSRNTSKSSKSSK SHRKRSSSKHRRSRSHRSKSRSKK SHRKRSSSKHRRSRSHRSKSRSK SHRKRSSSKHRRSRSHSKDKKKDTREKI KEKERVEKENDEKEREEREKEREKERGK KDRDKERERDKEKEREREKEREKERGK KDRDKERERDKEKEREREKEREKERGK KDRDKERERDKEKEREREKEREKERGK KDRDKERERDKEKEREREKEREKERGK KSTPPRSYNASRRSRSSSRSSRSSSSS PRTSTITIKRSSSPSPRSRKKKKITEKT HISERERERSTSMRKSSNDRDGKEKLEKNST S 375 1725 A 3192 415 101 AHSSHOTRAILQEFQWDIRHPPLSPNLALSG FFFNLKKSLRGTHFSSVKKITTLTWLNSQDP WFIFFYPISPDLQIPSSFRNGNDWYHHSQKC PDLDGAVVK FFFYPISPDLQIPSSFRNGNDWYHHSQKC PDLDGAVVK SQSAGITGVSTYTWFNHLITVLNSQDPS ASQSAGITGVSTYTWFNHLITVLAVISKRI 377 1727 A 3201 274 1285 KTGYTSRGSPLSPGSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDVASTSAS VSRRSSSVSLSSKKGTGSDGYDQVSLEDEE EFPHLPPPQPLLPRCSPPQRQIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTYSDQQP NRTNGDKPPKKYA*PSPDAKYNGH*QHSSP VTVRNQSFDSSLHGAGNGSRIQSCUPSQQL QHRVHSVGHFPSIRPLKATAYVSPTYQQS NMPLSNGLQLYSNTGIPTPNKAAASGINGSS ALPFSLAINGSNLPRSKLIQPVRSPLQPVRPL SISLSTLRDGNWRDGCY VGGYTESPPSVIRGDHLFALLSSETHOEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKKNPYTTROPQPR-VQIERALTGRWLLW PMLFPVAPPDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRAPYIIFGPFGTGKVTI LVPAIKQVVKHLFKAHLLAFSRDIRMVPEBLKPCN WDAAKGEYVPPAKKKLQFVKPLITTLITTAGR LVSAGPPIDHFTHIPIDEAGHCMEPESLVAIAG LLWSVKHLDFRSLALLAFSRDIRMVPEBLKPCN WDAAKGEYVPPAKKLLQFVRLATLLAFSRDIRMVPEBLKPCN WDAAKGEYVPPAKKKLLAFSRKA LLAFSSK, WASHSTILLIPROGLATION CHERGROSPSFTPREAATVTSTYCLPCAPTICIPE ETSSSFHSSPRRYTHEALATRAALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSYSLLWPPPLT-GSFKTKCTXCXIK	į	ŀ]]	1	ļ	NSQTTTADQLLEFFKQVGEVKFVRMAGDET
VREAQSFISAAIEPGWLHSTSLCNDFLGGF RR RMYRE*APCTICGFFHC LINDFLIPS ATTA K*FFPPRVWKEQ*PKRRRSRSHTRSKSRSSK SHRRKRSQSKHRSRSHRSRSRGVKRRRSKSSK SPHKKRSKSRRKSRSSHRSRSRGVKRKDTREKI KEKERVKEKDREKDEREKEREKEREKEKEK KEKERVKEKDREKDEREKEREKEKEKEK RDKEKEREQDKEKEREKEREKEKEKEK RDKEKEREQDKEKEREKORSKEIDEKRKDK KSRTPRSYNASRRSKSSSRSFRRRRSRSSSSS SPRTSKTIKRKSSSRS*PRSRNKKDKKREKER HISERERERSTSMRKSSNDRDGKEKLEKNST SPHRIKKSSKSPSPPSRNKKDKKLEKNST SPHRIKKSSKSPSPPSRNKKDKKLEKNST SPHRIKKSSKSPSPPSRNKKDKKLEKNST SPHRIKKSSKSPSPPSRNKKDKKLEKNST SPHRIKKSSKSPSPPSRNKKDKLEKNST SPHRIKKSSKSPSPPSRNKKDKKLEKNST SPHRIKKSSKSPSPPSRNKKDKLEKNST SPHRIKKSSKSSSPPSRNKKDKLEKNST SPHRIKKSSKSSSPPSRNKKDKLEKNST SPHRIKKSSKSSSSWDYR HVPPPANPVELDGFGHAGOGLIGDPPS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K K STGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMABLQESSLRQDVASTSAS VSRHSSSVSLSSKKGTGSDGSLDSTSELEDDSISM GYKLQDLTDVQIMABLQESSLRQDVASTSAS VSRHSSSVSLSSKKGTGSDGYDQVSLEDEE EFDHLPPPQPRLPRCSPPQRGIPHSQITESRIEC RRSPSSQYPSSNLYQGQVYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH*QHSSP VTVRNQSGSSSLHGAGNGISRIGSCDSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLOLYSNTGIPTRAAASGIMGRS ALPRPSLANGSNLPRSKIAQPVRSFLQPFKPL SSLSTLEDGNWRDGCY 378 1728 A 3202 112 1789 VPOVTESRPSVLRGDHLFALLSSETHQEDPIT YKGGFVIKVELDRVKLSFSMSLLSFTGWGF PFKVNFYTTROPQLRVQHRAPDIKPCON WDAKKGEYVFPAKKKLQPYRSLLTRGWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLSSP EQLQAMRHIVTGTRPAPUTHGPFTGTGKTVT LVPALKQVKHLPKAHLLACAPSNSGADLLC QRLRVHLPSSYRLLAPSSKKLTGRWLLW MDAKKGEYVFPAKKKLQPYVLITTLITAGR LVSAGPFIDHFTHIFIDEAGHCMEPESLVAIAG LMEWKETGDPGGQLVLAGDRPQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPGDPPQ FITKLLENYRSHPTLIDIPRQLYYTEGLQACA DVVDRERFCRWAGLIPRQGFPIHHGVMGKD ERGROSPSFTPREAATNTSLYTKGPGPOPQ FITKLLENYRSHPTLIDIPRQLYYTGELQACA DVVDRERFCRWAGLIPRQGFPIHHGVMGKD ERGROSPSFTPREAATNTSLYTKGPGPOPQ FITKLLENYRSHPTLIDIPRQLYYTEGLQACA DVVDRERFCRWAGLIPREATHTLIDIPRQLYYTEGLQACA DVVDRERFCRWAGLIPREATHTLIDIPRQLYYTEGLQACA DVVDRERFCRWAGLIPREATHTLIDIPRQLYYTEGLQACA DVVDRERFCRWAGLIPREATHTLIDIPRGLACHCMEPESLYALAG LMEWKETGDPGGGULAGDFGCGPVLRSPL TQKHGLGYSLERLLTYNSLYTKGPGDOPC FITSSFHSSPRPRPTAALNARALPEPLTPGD SNILRYWDGIRKPACLT	Į.	ļ		}		1	QPTRFAFVEFADQNSVPRALAFNGVMFGDRP
RMYRÉ*APCTICGTFHI-CLINWDL*LP-AYTA		i				•	LKINHSNNAIVKPPEMTPQAAAKELEEVMKR
	1				İ		
SHSRKRSQSKHRSSHNSRSKDKRKDTREKI		ĺ	1	[[[
SPHKKRSKSRERKSRSRSHSRDKRKDTREK KEKERVEKDREKERERKEKEREKKERKIN KEKERVEKDREKEREREKEKEREKEKERKIN KORDKEREKDREKDEKEREKEKEREKEKEREKIN KORDKEREKDOKEKEREKDRSKEIDEKRKDK KSRTPPRSYNASRSRSSSSSERRRRSSSSSSR PRTSKTIKRKSSRSPSPRSRNKDKKREKERD HISERRERESTSMRKSSNDRDGKEKLEKNST S	İ		}				
				ł			
	1		ł	ł			
RDKEKEKEQDKEKEREKDEKNEKDEK RKENK KSRTPPRSYNASRRSRSSSERRRRRSRSSSS PRTSKTIKKRSSRSPSPKSRNKKDKKREKERD HISERRERESTSMRKSSNDRDGKEKLEKNST S				Ì			
SSRTPPRSYNASRRSSSSRERRRRRSSSSSS RTSKTIKKSSRSPSPRSNKKDKKREKERD HISERREREESTSMRKSSNDRDGKEKLEKNST S S							
PRISKTIKRKSSRSPSRRNKKDKKREKERD		,	ļ	ĺ			
HISERRERSTSMRKSSNDRDGKEKLEKNST S							
S				i		•	1
375							
FFFPNLKKSLRGTHFSSVKKITTLTWLNSQDP WF/FFYP*SPDLQIPSSFRNGLNDWYHHSQKC PDLDGAYVKK	375	1725	A	3192	415	101	
WF/FYP*SPDLQIPSSFRNGLNDWYHHSQKC PDLDGAYVKK	5.5	1,23	**	1 31,72	113		
PDLDGAYVKK							
HVPPHPANFVPLLETGFLHAGQAGL\GDPPAS AS\QSAGITGVSHTWPKNHLIFYACL\VIRSKRI K 377 1727 A 3201 274 1285 KTGYTSR\GSPLSP\QSSIDSELSTSELEDDSISM GYKL\QDLTDV\QIMARL\QEESLR\QDYASTSAS VSRHSSSVSLSS\GKGTCSD\QEYDQ\YSL\QDE EFDHL\PPP\QPRL\PRCSPP\QR\GPHS\QTSSIT\RC RS\PSS\QYF\SN\YQ\QQ\YYS\PQA\QTP\QQ\P NRT\\GDK\PPK\YA\PS\PDAK\Y\\CH\PS\PV\YV\RN\S\SP\ST\SSL\GF\PS\RQ\Q\QT\PY\Q\QS\S\N\PB\S\RQ\Q\YS\F\Q\Q\T\PY\Q\GS\N\MPL\SN\GQ\L\YS\T\GT\PT\N\AA\AS\GIM\GRS AL\PR\PSL\AT\\QS\R\GT\PS\V\\GT\\RT\\QT\R\GT\PS\S\R\GT\\RT\\\QT\\\GT\\\GT\\\GT\\\GT\\\GT\							
HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS ASQSAGITGVSHTWPKNHLIFYACLV\TRSKRI K 377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQ\IMARL\QESEL\QDYASTSAS VSRHSSSVSLSSGKKGTCSDQEVDQ\YSLEDGE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RSFSSQ\YFPSNNYQQQ\YYSPQAQTPDQQP NRTNGDK\PPKYA*PSPDAKYNCH**\QHSSP VTVRNSQSFDSSLHGAGNGISRI\QSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTV\QGSS NMPLSNGLQL\YSNTGIPTPNKAAASGIMGRS ALPRSTLANGSNLPRSKI\AQPVRSFL\QPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRFSVLRGDHLFALLSSETH\QEDPIT YKGFVHK\VBLDR\VKLSFSMSLLSRFV\GWG* PFKVNFY\TFNRQPLR\V\QHRALELTGR\WLL\W PML\FPV\APRD\VPLIFSDVKL\KL\YDRSLESN\P EQ\QAMRHIVTGTTRP\AP\YII\FQPPGTGKT\T\ LVEAIK\QVVKHL\PKAHILACAPSNSGADLLC QRLR\YHL\PSSI\YRLL\APSRDIRM\VPEDIK\PC\N WDAK\GE\YVFPAKKKL\QETVL\ITTLITAGR LVS\Q\QFIDHFTHI\FIDEAGHCMEFESL\VAIAG LMEVKETGDPGGQ\VL\AGD\PRQ\GPPI\FIG\YMGKD ERGNS\PSFF\PDEAAT\TSYL\KL\LL\APSSKK GK\RL\S\PRS\PSV\S\U\S\YT\RQ\DP\YF\PGL\QACA DV\VD\RE\FCR\WG\DP\YF\G\PG\YD\PQ FITKLLR\N\YSSH\FTILDI\PN\QL\YY\EGEL\QACA DV\VD\RE\FCR\WG\DP\YF\RQ\G\PV\\RSP\L ERG\LD\R\G\S\S\F\Y\R\G\PY\F\T\Y\LL\L\APSS\K\G\GK\RL\S\F\Y\R\G\PY\F\T\Y\L\L\APSS\K\GK\R\S\F\Y\R\G\PY\F\T\Y\L\L\APSS\K\G\GK\R\S\F\Y\R\G\PY\F\T\Y\L\L\AP\S\K\G\GK\R\S\F\Y\R\G\PY\F\T\Y\L\L\AP\S\K\G\GK\R\S\F\Y\R\G\PY\F\T\Y\L\L\AP\S\K\G\GK\R\S\F\Y\R\G\PY\F\T\T\T\G\D\ S\R\R\W\R\S\F\Y\R\G\P\T\Y\T\S\H\S\S\K\G\F\T\Y\T\T\S\H\S S\R\R\W\R\S\F\T\Y\R\G\P\T\T\T\S\H\S S\R\R\W\R\S\F\T\Y\R\G\T\T\S\H\S\T\T\S\H\T\T\S\H\S\T\T\T\T\S\H\S S\R\R\W\R\S\F\T\Y\R\G\T\T\S\H\T\S\H\T\T\S\H\S\K\G\F\T\T\T\T\S\H\S S\R\R\W\R\S\F\T\Y\T\S\H\T\S\F\T\T\S\H\T\S\H\T\S\K\G\F\T\T\T\S\H\S S\R\R\W\R\T\S\F\T\T\S\H\T\S\H\T\S\H\T\T\S\H\T\S\H\T\T\S\H\T\S\H\T\T\T\S\H\T\S\H\T\T\S\H\T\S\H\T\T\S\H\T\S\H\T\T\T\T	376	1726	Α	3199	931	418	GV*WCDLGSPOPPPPGFKOFCLGRSSSWDYR
STOP							
377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PFKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISSYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRSVGVISSYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRPTPAAALNRARALPELTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI
GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDGE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFFSNNYQQQYYSPQAQIPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQLRVVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYPRSLESSP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAFSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISSYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRYPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							L -
VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVWELDRVKLSFSMSLLSRFVGWG** PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEALKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGGYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	377	1727	Α	3201	274	1285	
EFDHLPPPOPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQYYSPQAQTPDQQP NRTNGDK/PKKYA*PSPDAKYNCH**QHISSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSINGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPSSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRPPTPAALNRARALPEPLTPGD SNLRVWDGIRRPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK					-		GYKLQDLTDVQIMARLQEESLRQDYASTSAS
RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QHISSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKLQEYVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTTKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRPPTPAALNRARALPEPLTPGD SNLRVWDGIRRPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
NRTNGDK/PPKKYA*PSPDAKYNCH**QHSSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITIAGR LVSAQFPIDHFTHIPIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	·						
VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLANGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVLELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIPIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPSPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTTNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTILRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPFTPTAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
MMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			:				
ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKL\YDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
SSLSTLRDGNWRDGCY 378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRVQHRALELTGRWLLW PMLFPVAPRDVPLLPSDVKKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGULPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	· ·					Ï	
378 1728 A 3202 112 1789 VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKV\\(\subseteq \text{LSFSMSLLSRFVGWG*}\) PFKVNFY/TFNRQPLRV\\(\text{QHRALELTGRWLLW}\) PMLFPV\\(\text{PROVPLLPSDVKLKLYDRSLESNP}\) EQLQ\(\text{QAMRHIVTGTTRPAPYIIFGPPGTGKTVT}\) LVEAIK\(\text{QVVKHLPKAHILACAPSNSGADLLC}\) QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN \(\text{WDAKKGEYVFPAKKKLQEYRVLITTLITAGR}\) LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG \(\text{LMEVKETGDPGGQLVLAGDPRQLGPVLRSPLTQKHGLGYSLLERLLTYNSLYKKGPDGYDPQFITKLLRNYRSHPTILDIPNQLYYEGELQACA\) DVVDRERFCRWAG\(\text{LPROVEKTRYDEKLLAPSSKK}\) GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR \(\text{ERGRSPSFFNPEEAATVTSYLKLLLAPSSKK}\) GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR \(\text{ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLPETSSSFHSSPRPRTPAALNRARALPEPLTPGDSNLRV\(\text{WDGIRKPACLTNTSCHS}\) SNLRV\(\text{WDGIRKPACLTNTSCHS}\)							
YKGFVHKVELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLES\NP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	378	1728	Α	3202	112	1789	
PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
PMLFP\VAPRDVPLLPSDVKLKLYDRSLES\NP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPS\NSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK					. 1		· • • • • • • • • • • • • • • • • • • •
LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTIKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	1						
QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDREFFCR WAGYLPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTTKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK						!	
LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCTTKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK					İ	1	
LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAGILPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK							
TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK					J	1	
FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		i			l		, ,
DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			ļ		l		
EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK						İ	
GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK	([Ì	1			i	
ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK			1		I		
ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK					.		
SNLRVWDGIRKPACLTNTSCHS 379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK		ļ	j		ļ		
379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK				İ			
		1770				120	
	3/9	1729	A	3206	432	130	
							*LSTKEAXUSXPGKQIAXXKQGGKVETTTAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}	ļ	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i		ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			Ì	peptide		/=possible nucleotide deletion, \=possible
				sequence	<u> </u>	nucleotide insertion
						XKQSNNKGTRASSYXEPDAXEQWKFPHKKL OLPGXTHE
380	1730	A	3207	187	507	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS
						PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP
1	İ	ĺ				AXLLPGPGGGPGPVASLEARAQASSGVTPNG
201	1001	ļ. —	2005		0.40	GGRTYPYPTFSSGE
381	1731	Α	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/
						EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS
	1	İ			İ	KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI
]		J				WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP
		Į.				GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY
	1	ŀ				WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI
						KLQDMEKKANPSSLVLERREVEQQGFLHLGE
382	1732	<u> </u>	3238	256	38	HDGSLDLRSRRSVQEGNPRA
302	1/32	Α	3238	236	38	LLMIKVSSTCFSCHLHHHHHHHHHRHHQGHNS LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF
		l				LFSNACL
383	1733	A	3241	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKAD
						KVTMLWNKKATAVLVIASTDVDKTGASYYG
)]		ļ	EQTLHYIATNGESAVVQLPKNGPIYDVVWNS
						SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG
						PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK
	ļ					VWNVKNYKLISKPVASDSTYFAWCPDGEHIL
						TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP
1						NEEPKVATAYRPPALRNKPITNSKLHEEEPPO
						NMKPQSGNDKPLSKTALKNQRKHEAKKAAK
						QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP
						EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK
204	1724		2242			NQLEKIQKETALLQELEDLELGI
384	1734	Α	3242	3	678	RSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP
						LSOCARRVHGEKLRRPTFGPRHRGAGTAKMS
						ASLVRATVRAVSKRKLQPTRAALTLTPSAVN
						KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL
						EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL
						GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES
385	1735	Α	3243	3190	664	FNI VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL
1995	1733	^	3243	3170	VO*+	KEEEILPEPGSETPTVASEALAELLHGALLRR
						GPEMGYLPGPPLGPEGGEEETTTTITTTTVTT
1 1						TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL
						GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL
						VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT
		i 1				NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP
				Ì		PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA
						TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL
					ŀ	HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS
						DMDDVPERGLISDAQSLYVELLSETPANPLLL
					1	SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG
ļ .]			ALATFSCLPGYALEPPGPPNAIECVDPTEPHW
1						NDTEPACKAMCGGELSEPAGVVLSPDWPQS
			ļ	i	ļ i	YSPGQDCVWGVHVQEEKRILLQVEILNVREG
	[[ĺ		DML'I'LFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR
	l	-				NDTCPELPPPEWGWRTASHGDLIRGTVLTYO
}			Į			CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MTCADPGEIANGHRTASDAGFPVGSHVQYRC LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF
386	1736	A	3250	5725	3984	SNPLYEAGDTREYEVSI GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	Α	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQUEELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	A	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLIILGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RKRKPNKGESKEAIPDSKSLESSKEIPEQREN KQQQ
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS WICLSMVILTHSLKTFHRNWDWESEYTLFMS ALKVNKNNAKLWNNVGHALENEKNFERAL KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSSSVGPLRPGRPL WSEACAFL*AAAPQGPASPCCGLPSGFPRVW AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY STSFLTDSYLKYIGWTLHDKHREVRVKCVKA LKGLYGNRDLTARLELFTGRFKDWMVSMIV DREYSVAVEAVRLLILILKNMEGVLMDVDCE SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHKVFLFLLLPSLLMGYSE SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS GRAVALLHLIASGLTSIQTNTASSKPPIWGYL STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL TGAALAGSYPIWENENTLSWLPTFTYNFCLST PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI NILPPNQTILISVEASISSSPIRNKWALHLITLLT GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE DMHTSITSLQRQLDFLVGVILQNWRVLDLLT TEKGGTCIYLQEECCFCVNESGIVHLAVRRLH DRAAEL*HQVADSWWQGSSLLRWIPWVAPF LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRRELEIATSDNQE YYNRLCQEVTNRERNDQKMLADLDDLNRTK KYLEERLIELLRDKDALWQKSDALEFQQKLS AEERWLGDTEANHCLDCKREFSWMVRRHHC RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP EKIVLRALKDSRAGMPEQDKDPGVQENPDD QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF VVRTMCAVLGLVARQEDSGLRDHSVRVLISN HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
ĺ		i 1				GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ
						FFFF\IFLNIFLLAFSSPGSQPLLNSPPSFVCWSR
i i						GFMEMNGRGELVESLKRFCASTRLPPTPLLLF PEEEATNGREGLLRFSSWPFSIODVVOPLTLO
]		}				VQRTLVSVTVSDASWVSELL\WSLFVPFTVY
i i						QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE
						LG\OTGTRLTPA\DKAEHMKRORHPR\LRPOS
1 1						AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF
						CPHVAIGVFIPERPWPKTGCCKTLTIHLILL*G
						GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG
						PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL
						QERKQ\ALYEYARRFTERRAPGGLD
403	1753	A	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS
ĺ						GGASAGLASSPECACGRSHFTCAVSALGECT
						CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL
						DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD
						CPPRECEED
404	1754	Α	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG
						QDHVQNEEIYARVLDKFGSNFLSRDNADLGT
						AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS
						LLKGDLKGVKGDLKKPFDKAWKDYETKFAK
405	1755		2222	10	450	IEKEKREREWR
405	1755	Α	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA
						KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG
						HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIKWWE
						GKYTKPSQYNPNYMLELAHNDSVW
406	1756	Α	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV
				•	120	MCVLLWALSLLQSILEWMFCSFLFSDVDSDN
						WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC
						GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F
						LLYWIEKDLDDL
407	1757	Α	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID
						LNKVKTKTAAKYGLSAQPRLVDIJAAVPPQY
]						RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC
						PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP
						TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD
						KVEFIVMGGTFMALPEEYRDYFIRNLHDALS
1						GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC
			ĺ			MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESFHLAKDSGFKVVAHMMP
						DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP
			4			TLVIRGTGLYELWKSGRYKSYSPSDLVELVA
1					!	RILALVPPWTRVYRVQRDIPMPLVSSGVEHG
1						NLRELALARMKDLGIQCRDVRTREVGIQEIH
1			ļ			HKVRPYQVELVRRDYVANGGWETFLSYEDP
						DODILIGLERLERKCSEETFRFELGGGVSIVREL
1						HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA
I						ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL
ŀ			ĺ	-		QGPYMVKMLK
408	1758	Α	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK
						KGAKKKAV/DNINIGKTLVTRTQRTKIASDG
I						LKGRVFEESLADLQND\TDGYLLRVI*VAFTT
		ĺ	1			ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH
						DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	Α	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL
	,	ĺ				WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK
						WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL
410	1760	A	3339	127	1433	GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL
		^	3337	121	1433	WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLIP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342		2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSILLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEPLTERIKYD PIPVRWNSLEGRLLRGYEQHANDGKDYISRN *DLRSWTAADMAAQITKRKWEAEEFAEQIKA
						YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	Α	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
-	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	A	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID SEQ ID Met SEQ Predicted Predicted end Amino acid	d sequence (A=Alanine C=Cysteine,
	c Acid, E=Glutamic Acid,
	lanine, G=Glycine, H=Histidine,
	ne, K=Lysine, L=Leucine,
	nine, N=Asparagine, P=Proline,
	ine, R=Arginine, S=Serine,
	ne, V=Valine, W=Tryptophan,
	e, X=Unknown, *=Stop codon,
	nucleotide deletion, \=possible
sequence nucleotide	
	LVISKPVSKSPERLRKDIEVLSEDTD
	KKRKDVKKDTTDKSSKPOIKRGKR
	CLKTGSPGKKEEKAKNKESLCMEN
	EDEEETKAKMTPTKKYNGLEEKRK
	YSGFSEVAEKRIKLLNNSDERLONS
	DVWSSIQGQWPKKTLKELFSDSDTE
	IPAPEEGVAEESLQTVAEEESCSPSV
	VNVDSKPIEEKTVEVNDRKAEFPSS
	IPLPYLHLNRLHQSL*QKGSRQQSS
	APNQEEVRSIKSETDSTIEVDSVAGE
	RE*LASRF*CQCELKQ**SARTRTS*
	KSERCSGRRKFIKKAEKKP*SNSGK
OOKEGK	RSERCSCIAGO INICALIGIA SINSCIA
	HIGQAGIQIGDACWELYCLEHGIQP
1 . 1	TOODOLENAKMEHTNASFDTFFCE
	VPRALFVDLEPTVIDGIR
<u>' </u>	WSSVLVTQARVQWRDLGSPQPLP
	CLSLPSSWDYRHPSPRPVNF/HVFLV
	VGQAGLELLTSGDLPALASQSARIT
	QPRGHFH
, -: -: -: -: -: -: -: -: -: -: -: -: -: -: -: -: -: -: -:	SCWQELGLGPWGGDWRVEQVGAS
	EVCSIRFLFTAVSLLSLFLSAFWLGL
	ENEPKEMLTLSEYHERVRSQGQQL
	LDKLHKEVSTVRAANSERVAKLVF
	VRKPDYALSSVGASIDLQKTSHDY
	YFWNRFSFWNYARPPTVILEPHVFP
	EGDQGQVVIQLPGRVQLSDITLQHP
	GGANSAPRDFAVFFLLSFFTHQGLQ
	/SLGKFTFDVEKSEIQTFHLQNDPPA
	QILSNWGHPRFTCLYRVRAHGVRT
SEGAEGS	
	IGALVVFKRP*ATTGSDPGPKRGMN
	IRSPESGKGEPGTARDYTPMGRPPP
	GPLPGSLAIAPHSPEPHPWEQQPPRG
	GWLGSAT/RVRRPHNHP/RGH/HSP
VDIAGAI	PASPGPDVCE
	SSVGPAVSLRQRQQDGAVKESGR/
I I I I I I I I I I I I I I I I I I I	SRAAAAMAPIKVGDAJPAVEVFEG
	VLAELFKGKKGVLFGVPGAFTPGCS
	VEQAEALKAKGVQVVACLSVNDA
	GRAHKAEGKVRLLADPTGAFGKET
	LVSIFGNRRLKRFSMVVQDGIVKA
	TGLTCSLAPNIISQL
	GTRVLGSTTAAVFLSVEDDNDNAPQ
	VQVREDVTPGAPVLRVTASDRDKG
1 1 1 1 1 1 1 1	YSIMSGNARGQFYLDAQTGALDVV
	TKEYTLRVRAQDGGRPPLSNVSGL
	DINDNAPIFVSTPFQATVLESVPLGY
	AIDADAGDNARLEYRLAGVGHDFP
	GWISVAAELDREEVDFYSFGVEAR
	TASASVSVTALDVNDNNPTFTQPE
	DAAVGTSVVTVSAVDRDAHSVITY
	RNRFSITSQSGGGLVSLALPLDYKLE
RQYVLAY	VTASDGTRQDTAQIVVNVTDANTH
RPVFQSS	HYTVNVNEDRPAGTTVVLISATDE
	RITYFMEDSIPQFRIDADTGAVITQA
, , , , , , , , , , , , , , , , , , , ,	VSYTLAITARDNGIPQKSDTTYLEI
	DNAPQFLRDSYQGSVYEDVPPFTSV

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	
nucl-	peptide	1100	in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine. G=Glycine. H=Histidine.
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchac		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ucite			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1	1	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		İ		peptide	sequence	/=possible nucleotide deletion, \=possible
	[sequence		nucleotide insertion
——				sequence		
						VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
	1					GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
1						VFVEENSPIGLAVARVTATDPDEGTNAQIMY
						QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
1						YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
						LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
i i						DISDSLTYSFERGNELSLVLLNASTGELKLSR
1						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
						TITTDEMLTHSITLRLEDMSPERFLSPLLGLFIQ
						AVAATLATPPDHVVVFNVQRDTDAPGGHILN
1						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
						LLTAISAQRVLPFDDNICLREPCENYMRCVSV
1 1	l					LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1						TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
1						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
1						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
						YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
						TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
						QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
						VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
1 1	ĺ			{		VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
				i		ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
1 1						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
!			- 1	ļ		LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
						GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
1						QASSLRLEPGRANDGDWHHAQLALGAIGGP
		1	-	j		GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
1 1						GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
						SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
į						CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
			- 1			EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
			İ	1		RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
		- }				NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
			J			SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
1	l l		- 1			AEVTTNGCEVNYDSCPRAŒAGIWWPRTRFG
	į	Ī	ļ	l		LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
		1]	}		NCTSITFSELKGFAERLQRNESGLDSGRSQQL
						ALLLRNATQHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL
		ļ	l			DTANKRHWELIOOTEGGTAWLLOHYEAYAS
1	ľ	ì	i			
[1	ŀ		ļ	ļ		ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR
		İ	ŀ	ļ		ETPPVVRPAGPGEAQEPEELARRORRHPELSO
]]	j	J	j	J		GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
	ļ	l	l	ļ		
j [ŀ	1		1		RPIINTPVVSISVHDDEELLPRALDKPVTVQFR
1	[l			LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCOCNHMTSFAVLMDVSRRE
1 1	l	l	ł	ļ	ŀ	NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
		l		ł		RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
; l	ŀ	j	İ			DLPFACTVIAILLHFLYLCTFSWALLEALHLY
j l	1	l	l	ļ		RALTEVRDVNTGPMRFYYMLGWGVPAFITG
	l	(ſ	· [
1		l	!	ļ		LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG
	ŀ		-	1		PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
	ł	l	}	ł		-
		l		1		FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
. 		I	Į.	1		LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
	ł		ľ	ŀ		RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
j 1	1	l		1	}	EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE
						TITES TUTE FOR PARTIES DE LA STESSE DE LA ST

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP EERLRENGDALSREGSLGPLPGSSAQPHKGIL KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
425	1775		3429	155	1417	GTVDEDSSGSEFLFFNFLH GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS RAASVREAEDAPLQPASIHPVSQGSRGPEGSL GSAECLPGDPLGARRATRAHSPVPGPPPSLPA AGTAVKRGLQPG*GA/GATSTPGTGAATGGL CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGMVILHRISAADPARAPGPDS NLQSALQQPATGCSEPAAVYSPPIGLWGA**P EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS
426	1776	A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTG\SHDFAINFDPENPECEGK RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVFI\EISGVIASYRRCLPQIQLYGPTNVAP IINRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP
427	1777	A	3446 .	79	9748	GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPP

No of body and the control of the co	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
nuclecide sequence where the contemporaries of the contemporaries							
uence USSN 914 14 15 15 15 15 15 15 15 15			nod .		•		1
uence 94496 Orresponding of the first and one of peptide residue of peptide sequence Poptide sequence							
uence 1914 ng no first amino acid residue of peptide pept	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence 1914 ng no first mino acid residue of peptide peptide peptide peptide peptide sequence P-Threonine, V-Valine, W-Tyrpophan, Y-Tyrosine, X-Unknown, **-Stop codon, P-possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide detailor, *possible nucleotide, *possibl	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per				3			
residue of peptide sequence Y=Tyrosine, X=Unknown, **Siop podon, /*possible nucleotide delicin, *possible nucleotide diction, *possible nucleotide diction, *possible nucleotide insertion TTFYPEEGYVSDILAYIDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTRTI.TGNTFSI. ADCIPLILRRII.KDESSYTCKI.ACTAYRCVM SLCSSSYSEI.GL.QLIDVI.TI.RNSSYW.J.WITEL LETLARIDPRI.VSTELAKARDH.RRGAHFYTGI. LRLQERVI.NNVVHILLGDEDBRYRHVAAASI. RILVPELEYKCDQG.QADPVVAVARDQSSYVI. KILLMHETQPSIFSVSTITRIYRGYNLLPSITD VTMENNI.SRVIAAVSBUELTSTTRAI.TGGCC ALCILSTAPPVCIWSI.GWHCGVPPL.SADESR SCCTVGMATMILTI.LSSAWFDLDASAHQDAL LLAGNILAASAPSI.RSSWASEEANPAATK QEEVWPALGDRALLVPWGQLFSHLLKVNIC AHVI.DDDVAPCPAIKAAI.PSI.TDFSI.SPIRRK GEKEPDGGASVIPLSKGSSRASARGSDTS GPVTTSKSSSI.GSFYHLPSYLK.HDVI.KATHA NYKVTLDLONSTERFGGFI.RSALDVI.SQUIEL ATLQDIGKCVEELIGYI.KSCFS.REPMMATVC QQLLKTLIGTINAASOFGOI.SSNPSKS.GGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMYQAEGEMDTSGWFDU.CQVSTQLKT NILTSYTINRABOKNAIBHNHRI.FEPL.VIKALKQ QYTITCVQLQKQVLJLAQLVQLEVYCQL SYDYSTRYSTRYSINGIGPRINGLOCORIANSCR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LETOKEVVSMIL.BLI.QVPQVCPRESEAIIPNIFF FLVLLSYREYTSKOIGPRINGLOCORIANSCR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LETOKEVVSMIL.BLI.QVPQVCPRESEAIIPNIFF FLVLLSYREYTSKOIGPRINGLOCORIANSCR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LETOKEVVSMIL.BLI.QVPQVCPRESEAIIPNIFF FLVLLSYREYTSKOIGPRINGLOCORIANSCR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LETOKEVVSMIL.BLI.QVPQVCPRESEAIIPNIFF FLVLLSYREYTSKOIGPRINGLOCORIANSCR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LETOKEVVSMIL.BLI.QVPQVCPRESEAIIPNIFF FLVLLSYREYTSKOUGESTSTILL SHEEPEV STALLEV SURGESTSTILLS GAMFRITAAATRI.FRSDCCGGSFYTLDSILNER GREFRITA SHEEPEV STALLEV S]			
peptide sequence nucleotide insertion TEYPEEGYVSDILAYIDHOPPQVRGATALLC GTLICSLERSERPHYDDWMGTRTLTONTESL ADCIPLLEKTI.KDESSVTCKLACTAVRNCVM SLCSSYSYEGLQ.LQLIDVITLRNSSYW.NYTEL LETLAEIDFRLVSTLEAKAENLHRGAHHYTGL LELGERVLNNVVHLLGDEDRVRVHAAASL RLVPKLFYKCDQQ.ADPVVAVARDQSSVYL. KLLMHETQPSHIRSVSTTIRINGQYMLLSTID VTMENNI.SRVAAVSHELITSTTRALTFGCCE ALCLESTAPPQVINSL GWHCCYPPLSASDESR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSR KSCTVGMATAMI.TLLSSAWFBLD.SANDSAT LAGNILASASPKISTAWFSQUFFECKLERY ANVENTAMI.ASAPKISTAMI.TSANDSAT GREVEPPALGRALYSSAMFSANDSATS GVYTTSKSSAGNAARROSDTIS ANVENTAMI.TSANDSAGNAI.TSANDATA WYKVTLDLONSTERFGGFLRSALDVILSQILEL ATLQDIOKCVSEELOYLKSGFSREPMMATVC VQULKTLTGTIALASQFSTD.SSANTSAGNA SLRNMVQAEGENDTS ONF DVLQKVSTOLKT NITSYTKNRADKNAIBHIRLFEPL VIKALKQ VTTITCVQLQKQVLDLAQLVQLRVNYCLL BSDQVFIGFVLRQFEYEWGQFRESEALINNIFF FLYLLSYRYSKGNGTONFNOLGOMASGR KAVTHAIPALQPVINDLFVLRGTNKADAGKE LETYKEVVVSMLLALIQVINQLEMPILVLQQ CHKENEDKWKRLSRQIADILERMLAKQQMH DSHEALGVVVSMLLALIQVINQLEMPILVLQQ CHKENEDKWKRLSRQIADILERMLAKQQMH DSHEALGVVTNIFLELAPSSRRVDMLLSSMF VTPNTHAASVSTVQLWINGGLALIEVLISQSTEID IVLSRQIESTSPYLLGLVYNIRLRCDDSTSTILE BHSEGRQIKNLPEFTSRFLLQLVGILLINNITYRW ARVQQTFRKHSLSSTKLLSRQMSGEEDSDLA AKLGMCREVKROALIFCDVYCQHLDSS HLTWLINNIGDLISLSTEPPLAGLINGEY LQSSGLAQRHQRYSLLFCDVYCQHLDSS HLTWLINNIGDLISLSTEPPLAGLINGEY LQSSGLAQRHQRYSLLFCDVYCQHLDSS HLTWLINNIGDLISLSTEPPLAGLINGEY LQSSGLAQRHQRYSLLFCDVYCQHLDSS HLTWLINNIGDLISLSTEPPLAGLINGEY LQSSGLAQRHQRYSLLFCDVYCQHLDSS GCWTRSDSALLEGABLVNRIPAEDMNAFM MNSEFINLSLAPCLLSCHNSEGGGGSALFEA ARVTLARVSGTYQLWHVFLEAVAQQPG GELLSFERMTTPKAKSLECTVSPKDWVYHLVX SQCWTRSDSALLEGABLVNRIPAEDMNAFM MNSEFINLSARTLSLTAPLDGLOPKYHLEAVAQQPG GELLSFERMTTPKALSGREEEVDPHTONNINNIC SWCREEDFELDERTONNLAVAGATSLVSLAM VPV							, , , , , , , , , , , , , , , , , , , ,
micleotide insertion TTYPEEQTVSDILAYIDHGDPQVRGATAILC GTILCSILSRSRPIVGDWMGTRTI.TGNTFSI. ACCIPLLEKTLKDESSYVCKLACTAVRCVM SLCSSSYSELGLQLIDVITLRNSSYW.NYTEL LETLAEDFRUNSTLEAKABNLHRGAHFYTGL LETLAEDFRUNSTLEAKABNLHRGAHFYTGL LEKLQERVLNNVVHLLGDEDBYNRHVAAASL RILVFELLFYKCDQGOADPYVAVARDQSSVYL KILMHETQPPSHFSVSTITRIYRGYNLLPSITD VTMENNI.SRVIAAVSBLEITSTTRALTFGCCE ALCLISTAPPVCIWSL.GWFCGYPPLSASDESR RSCTVGMATMILTILSSAWFFLDLSSAMPDAL HAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGPRALVPWEQLFSHLLKVNIC AHVUDDVAPPAIKAALPSILTSPFSLSPIRKR GEKEPGGASVPLSKGSESAASRGSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVI.KATHA NYKVTLDLONSTERFGGFLRSALDVI.SQUEL ATLODIGKCVERLIGYUKSGFSREPMMATVC VQULKTLFGTINLASGFOEDISSNPSKSGGRA QRLGSSSVRPGLHYCFMAPYTHFTQALADA SLRNMVQAGEOBNITSGWFDVLQKVSTQLKT NILTSVTKNRADKNAIRNHRILFEPLVIKALKQ YTTTTCVQLQKQVLULLAQLVQLEVNYCLL DSDQVFIGFVLKQFEVEVGGPRESEAIIPNIFF FLVLLSYRFYHSKGIGFFRIQLCOTIONASGR KAVTHAIPALQPIVIDLEVLRGTNEADAGKE LTOKKEVVSMILAILQUVJQHVQLEMFILVLQQ CHKENEDK WRKLSRQIADILRPALAKQQMHI DSHBALGVVTNTFEILAFSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAIRVLISQSTED JVSRIGGEJSFPPLLSCTVINRLRDGDSTSTILE BHSEGRQIKNLPETSRRPLLQUIGILLEDIVT RQLEVEMSQQIATIVPALWENSTTYLL RABSGRGVINLPETSRRPLLQUIGILLEDIVT RQLEVEMSQQIATIVPALWENSTTYLE BHSEGRQIKNLPETSRRPLLQUIGILLEDIVT RQLEVEMSQQIATIVPALWENSTTYLE BHSEGRQIKNLPETSRRPLLQUIGILLEDIVT RQLEVEMSRQQIATIVPALWENSTTYLE BHSEGRQIKNLPETSRRPLLQUIGILLEDIVT RQLEVEMSRQQIATIVPALWENSTTYLE BHSEGRQIKNLPETSRRPLLQUIGILEDIVT RQLEVEMSRQQIATIVPALWENSTYLQLUSGSTED JVSRIGGEJSFPPLLSCTLYDRICMDVIHVS QARVENSRQQIATIVPALWENSTYPTLKTLQLEGI HLSQSGAVLTLYVDRLLCTPRFVCARMVDIL ARRUMANDLILLSHPPVQOFISAVHRN ASCIPLIFORYCQGLANDSSTTLE BHSEGRQKIKNLPETSRRPLLANDLL ARRUMANDLILLSHPPVQDFISAVHRN ASCIPLIFICATIVPALWANDLL ARRUMANDLILLSHPPVQDFISAVHRN SQCWTRSDSALLEGABLVNRIPAEDMNAFM MISEPILSLAPCLS GMSEIGGGRSALFEA AREUTLARVSGTQQLJLAVHHYFQPELPAEP AYWSKLENTHAKSGEEDEDPDFTONPKYT TAACEMVAEMVSELQSVLALGIKRNSGVPA BLITTARVSGTQQQLAVHHYFQPELPAEP AYWSKLENSRQALFERVFELQGRWYNKLICALLPRINGLY LVVVSKLPSHLHHLPPEKEDDIVKFVVATLEAL WHILBBEQFILSDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLUYCHFILEAVAQQPG GESPPEEDTERTQNNLAVAGLIGCSM				'		sequence	
TTEYPEEGYVSDILANTIDHODPQVRGATAILC GTLICSLERSREPHYGDWMGTRIT. IONTES! ADCIPLLRKTI.KDESSVTCKLACTAVRNCVM SLCSSYSEGLGQ.LIDVUT.IRNSSYW.NTREL LETLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLAEDFRLVSFLE.AKAENLHRGAHHYTGL LLAEDFRLVSFLE.AKAENLHRGAHHYTGLE LLAEDFRLVSFLE.AKAENLHRGAHHYTGLE LLAEDFRLVSFLE.AKAENLHRGAHHYTGLE LLAEDFRLVSFLE.AKAENLHRGAHHYTGLE LLAEDFRLVSFLE.AKAENLHRGAHHYTGLE.STE VITMENNI.SRVAAVBELTISTTRALTFGCCE ALCLESTAEPVCWING.GWHCCYPPL.SASDESR KSCTVGMATMILTLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR KSCTVGMATMILTLLSSAWFILD.SADESR AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE AFTELDAARDAATE ATTUCKTURAENCH ATTUCKTUR	1				peptide	•	/=possible nucleotide deletion, \=possible
GTLICSILSRSRFIVGDWMGTRTLTONTESI ADCIPILERITADESISYTCKLACTAYCRNCVM SLCSSYSEIGLQLIDVLTLRNSSYWLRTEIL LETLAEDPRILVSTELAKARANLRRGAHHYTGL LELAGERVLNNVVHLLGDEDPRVRHVAAASL RIVVPELEFYKCDQG,ADPPVAVAABQSSYYL KLLMHETOPPSHESVSTTRIYRGYNLLPSITD VTMENNLSRVLAAVSHELITSTTRALTFGCYELLPSITD VTMENNLSRVLAAVSHELITSTTRALTFGCYELLPSITD VTMENNLSRVLAAVSHELITSTTRALTFGCYELLPSITD VTMENNLSRVLAAVSHELITSTRALTFGCYELLPSITD VTMENNLSRVLAAVSHELITSTRALTFGCYELLPSITD VTMENNLSRVLAAVSHELITSTRALTFGCYELLPSITO VTMENNLSRVLAAVSHELITSTRALTFGCYELLPSITO VTMENNLSRVLAAVSHELITSTRALTFGCYELLPSITO VTMENNLSRVLAAVSHELITSTRALTFGCYELSAGER KSCTVGMATMILTILSSAWFPLDLSAHQAAL LLAGNLAASAFKSITSSWASEEARAPAATK QEEWWALGDRALPSTRWEGYESHRAATK GKEKEFGEQASVPLSFKKGSES-SAASRQSDTS GPVTTSKKSSLGSFYHPLSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGCVEELLGYKSCSFERPMATVC VQULKTLFGTTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLKSCSFERPMATVC VQULKTLFGTTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLSCSFERPMATVC VQULKTLFGTTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLSCSFERPMATVC VQULKTLFGTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLSCSFERPMATVC VQULKTLFGTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLSCSFERPMATVC VQULKTLFGTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYKLSCSFERPMATVC VQULKTLFGTNLASGFDGLSSNPSKSQGA QRLGSSSYPEQLHYTCYLAGATACHATA LLTSTRYNCHACHATACHATACHATACHATACHATACHATACHATAC	1				sequence		nucleotide insertion
GTLICSILSRSRPIVGDWMGTRTLTIOTITSI. ADCIPILERIT. IADESSYTCKI. ACTEVIRCYM SLCSSYSELGLOLIDVI.TI.RNSYWLVRTEL. LETLAEDFRILVSTELGA. ADPIVAVAADASI. RLVPKLETYKCDQG. ADPVAVAADQSSYYL. KILMHETOPPSHESYSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTIFGCYS. ALCLLSTAPPVCTWSLGWHCGYPPLSASDESR KSCTVGMATHLTILSSAWFPLDLSAHQDAL. LLAGOLLATASAPKSIRSSWASEEANPAATK QEEWWPALGDRALPSTITPOLSSHOPATA GEEWWPALGDRALPSTITPOLSSHOPATA GEEWWPALGDRALPSTITRESPIRKS GKEEFGEQASVFLSFKKGSEASARGSDTS GFVTTSKSSSLGSFYHLFSYLKLHOVI.KATHA NYKYTLDLQNSTEKFGGFLRSALDVLSQILEL. ATLQDIGCVEELGLYKSCSFSEPMATVC VQQLLKTLFGTINLASQFDGLSSNPSKSQGA QRLGSSSYPECHTYLCYMAPTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYTKINRADKNAHMHRIFEFPLYKLALDV VTTTTCVQLQKQVLDLLAQL.VQLRVNYCLL DSDQVFIGFVLQGFEVEVGQFRESEALINHFF FLVLLSYBRYHSKQHIGPKHQLCDGIMASGR KAVTHAPHALQPYIVBLDFVLGTNIFADGSKE LETQKEVVVSMLLRLQVHQVLEMFILVLQQ CHENEDENWKRISRGADILBMALKQMHH DSHEALGVLNTLFEILAFSSLRPVDMLLRSMF VTPTIMASVSTVQLWISGILLBELVLSQSSTED IVLSRIGGELSFSPYLISCTVINRLRGGSTSTILE BHSEGKQIKNLPEETFSRFLLQLVGILLEDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFFSRFLLQLVGILLEDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRFLLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLLBLDIVY KQLKVEMSEQGHTFYCOELGTLLMCLBIEKS GMFRRITAAATSLFSRELLQVGNLEDINL ARSMITTIPBALVLLWCQULLLVMHTDYRWW ARVQTTFRHAAATSLFSRELLQVGGGNSALFAA ARLGMCOREVERGAAULTLYCDVCQNLIDSE HT-ULVINHIDDISLS HEPPVOPFISH VIRNY ARSGIFLAARQLSTRAAATALFTSRELLQVGGRESDSLA ARLGMCOREVERGAAULTLYCDVCQNLIDSE HT-ULVINHIDDISLS HEPPVOPFISH VIRNY TAACEAWABMVSLQGAVGRAVFRIVAALQFT LVVSKLISHLHLIPPEKRDIVKFVAATLALQ WSVSSTEPVTHACSLIVCVHFLEAVAVQPC GESPPEEDTERTQNNLAVAATSLVLSAMT VPAAGNSCLGGGRAVTLYVORTLLGALVGRIPTIONKYT TAACEAWABMVSLCQGAMYSKEMAATSLALQFU VPLASHAPAVSCLCQGPRINKPYLALLGAVVGU UVSSLAFBRITTHAASGEEEDEDYPTIVNKYT TAACEAWABMVSLCQGAMYSKE							TTEYPEEOYVSDILNYIDHGDPOVRGATAILC
ADCIPLIERTIKDESSYTCKLACTAVRNOWN SCSSYSEGUGUGUNUTIARNSSYWLYTEL LETLAEDFRLVSFLEAKABRIHRGAHHYTGL LELGERVLNNOWHLGDEDRYNWAAASL RLVPKLFYKCDGGADPVVAVARDGSSVYL KLLMHETOPPSHFSVSTITHYRGYNLLPSITD VTMENNIS RVIAVSHELITSTRALFGCCE ALCLLSTAFPVCWSLGWHCCVPFLSASDESR KSCTVGMATMITLISSAWFRILESAHDGAL ILAGNILAASAPKSLRSSWASEEBANPAATK GEEWPALGPRALFWAEQLFSHLLKVINIC AHVLDDVAPGARKAALFSLTNPFSLSPIRRK GEKEPFGCASVPLSKGSESAASRGSDTS GPVTTSKSSSLGSFYHLPSYLKHDVLKATHAA NKVTLDLQNSTERFGGFRSALDVLSQUEL ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQULKTLGTINLASGYPGLSSNPSKSGGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRIMWQAGEORDTS GWPDLQKVSTQLKT NITSYTKNRADKNAINHIRLFEPLVIKALKQ YTTTCVQLQKQVLDLLAQLVQLEVNYCLL DSDQVFIGFVLKQFEVEVGQPESSEAIPNIFF FLVLLSYREYHSKGIGFKIQLCDGIMASGR KAVTHAIPALQPVIÐILFVLRGTNKADAGKE LETOKKSVVSMLLLIQUVIGLENVCLL DSDQVFIGFVLKQFEVEVGQPESSEAIPNIFF FLVLLSYREYHSKGIGFKIGLCDGIMASGR KAVTHAIPALQPVIÐILFVLRGTNKADAGKE LETOKKSVVSMLLALIQUVIGVELBURHLVLOQ CHKENEDKWKRLSRQIADILPRILAKQOMH DSHBALGVLNTLFELAPSSLRPVDMLLRSMF VTPTIMASVSTVQL WISGILAILEVLISGSFED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE BHSEGKQIKNLPBETFSRFLLQLVGILLEIDIVT KQLKVEMSEQQHTFYCQELGTLLMCLLIIFKS GMFRRTTAAATRLFSRFLLQLVGILLEIDIVT KQLKVEMSEQGHTFYCQELGTLLMCLLIIFKS GMFRRTTAAATRLFSRFLLQLVGILLEIDIVT KQLKVEMSEQGHTFYCQELGTLLMCLLIIFKS GMFRRTTAAATRLFSRFLLQLVGILLEIDIVT KQLKVEMSEQGHTFYCQELGFLLMCLLIIFKS GMFRRTTAAATRLFSRFLLQLVGRLLEIDIVT KQLKVEMSEQGHTFYCQELGFLLMCLLIIFKS GMFRRTTHAAATRLFSRFLLQLVGRLLEIDIVT KQLKVEMSEQGHTFYCQELGFLLMCLLIIFKS GMFRRTTHAATRLFSRFLLLGVMGGEEDSDLA AKLGMCKREIVRRGAALLFCDVYCQNLHDSE HTUTUVNIQDLISLSHPPVOPDFLSAVHEN AASGLFIQAIGSRCENLSTTTMLKKTLQCLEGI HLSGSGAVLTLYVDRLLISPFRVLDFTGSLSSFLAVFINS AASGLFIQAIGSRCENLSTTTMLKKTLQCLEGI HLSGSGAVLTLTYVDRLLIFFRYKANTVORLS VYNSTEFFVTHACSLIVCVHFLLARALQOY LVVSKLIDLFFRSTKNOTHANAM MISEFNLSLLAPCLS GMSEISGGGKSALFEA AREVITLARVSGTVOQLYAHHVFOPSLAPE AAYWSKLNDLFGBAALYOSLYPLVWKLG WSPKPGGDFGTAFFEERVEFLQGKEVKKEFTYR TAACEMVABMVSELQGAMYSKENLATSHLVVAGL UNSKLIPPLYNGTRISKULLGTRFRYKLUKGL VPNSKLYBBEDFEDERFRONLAVVAGPG GESPPEEDTERTQNNLAVALTSLULSAMT VPNAGMPAVSCLEQGPRINKLILLGUNFREKLUKGUNF	1						1
SLCSSYSELGLQLIDVLTLRNSSYWLNTEL LETLAEDPRIVSELEAKAENLINGAHIYTGL LKLQERVLNNVVHILLGDEDPRYRHYAASL RLVPKIFYKCDQQADPVVAVADQSSVYL KLLMHETQPRSHESVSTITRIYGCYSLLLSFITT VTMERNIS.RVLAVASHELITSTRALTFGCCE ALCILSTAPPVCIWSLGWHGVPPLSADESR KSCTVGMATMITTLSSAWFILDSAHDDAL LLAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPAVEQLFSHLLKVINIC ARVLDDVAPGFARAALPSLTMPSLSHIPSLSHIPSL GKEKPGGQASVPLSKKSGSASAASROSTS GVTTSKSSSLGSYFHLFSYLKLINDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILE ATLQDIKCVEELIGVYLSCSSTERPMATVC VQQLLKTLFGTNLASQFPGLSNNSSSQGA QRLGSSSVRGGLYHCSTSRSFMATVC VQQLLKTLFGTNLASQFPGLSNNSSSGGA QRLGSSSVRGGLYHCSTSRSFMATVC VQQLKTLFGTNLASQFPGLSNNSSSGGA AQRLGSSSVRGGLYHCSTSRSFMATVC VQQLKTLFGTNLASQFPGLSNNSSSGGA QRLGSSSVRGGLYHCSTSRSFMATVC VQQLKTLFGTNLASQFPGLSNNSSSGGA QRLGSSSVRGGLYHCSMATHFTQALADA SLRNMVQAGENDTSGWFDVLQKVSTQLKT NLTSVTKNRAANAHNBIRLFEPLYKALKQ YTTTTCVQLQKQVLDLAQLVQLRVNYCLL DSDQVFIGFVLXQFEYEVGGGRESSAHIPNIF FLVLLSYERYHSKQIIGPKIQLCDGIMASGA KAVTHAPALQPIVHDLFVLGGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CKKENEDKWKLSKQADDILINGAQOMH DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNITMASVSTVQLWSGILALRVISSSTED IVLSRIQELSFSPYLISCTVINRLRGDGTSTIL BISGGGKINLPETFSRFLLQLVISSSTED IVLSRIQELSFSSPYLISCTVINRLRGDGTSTLUSLNLR ARSMITHPALVLLWCQILLLVNHTDVRWW ARVQOTPKRINLSSTKLSPONGLLLFONTYR QLKVEMSEQQHTFYCQELGTLLMCLLIBERS GMFRITAATRILFSDGCGGSFYLLDSLNLR ARSMITHAALVLLWCQILLLVNHTDVRWW ARSMITHAALVLLWCQILLLVNHTDVRWW ARSMITHAALVLLWCQILLLVNHTDVRWW ARSMITHAALVLLWCQULLLVNHTDVRWW ARSMITHAALVLLWCQULLLVNHTDVRWW ARSMITHAALVLLWCQULLLVNHTDVRWW ARSMITHAALVLLWCQULLLVNHTLXKTQCLSGI HLSQSGAVTLTVYDRLLCTFFRYLDONLSPS LQSGAVTLTVYDRLLCTFFRYLDONLSPS LQSGAVTLTVYDRLLCTFRYLKYTQCLSGT LAGRAGAALTSLAPCLLAQPIC LVVSKLPSHLLBCDGHVSLLTRAALQCY LVVSKLPSHLBLDGOHVSLETVSPDKDWYMLVK SCCWTRSSALLBCAALVNRTRACHONLSVRY ARCHITARRAGGRAALYSLILBRENGENDAFA HLTBLLRNIBLSAARPSLUCARPHVLWKLG WSPYRSGDFGTAFFERPELQEKEVFKETYTY TAACEMVARMYSLCROPTIVALACQATISLVLSAMT VPVAGADAALYSLTRAICREELLCARPUPLURFER LITELRRNIBLAARPPLVNSTTKCPVLYWKLG WSPKEGDFGTAFFERWATLGLARCAPPLYWFUR QESSPEEDTBROTINVLAVQAITSLVLSAMT VPVAGARAYSLCGCQPRINCKLALAGREN	1						
LETT.AEIDER.U.ST.EAKAEN.HR.GAHH.YTG. LKL.QERVLNNVVHLLGDEDR.YNHAAASL IRLVPKLFYKCDQQADPVVAVARDQSSVJ. LKLLMHETOPPSHERSVSTITRIYR.GYNLLPSITD VTMENNI.SRVIAAVSHELITSTTRALTFGCCE ALCLLSTAPPVCINSLEWHGGYPSLASDESR KSCTVGMATMILTLISSAWFPLDLSAHQDAL ILAGNILAASAPKSLRSSWASEEBANPAATK QEEVPPALGDRAKAALPSLTNPPSLSPIRK (REKEPGGASYPSPKKSGEASASRQSDTS GFVTTSKSSSLGSFYHLPSYLKLHDVI.KATHAA NYKVTIDLQNSTEKFGGFI.RSALVSQUEL ATLQDIGKCVEELGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFGGLSSNSGGAA QRLGSSSVRPGL.YHYCFMAPYTHFTQALADA SLRNNVQAGEDBITSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHIRLFEPL.VKALKQ YTTTTCVQLQKQVLDLLAQLVQLWSVCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFWCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFMLLAVUCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFMCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFMCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFMCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGBFKIQLCWFMCLL DSDQVFIGFVLKQFEYEEVGGPRESEAHPNIFF FLVLLSYERYHSKQIGFKIQFIQLOCH CKENEDEWKKLSRQIADILPMLASCOMH DSHBALGYLNTLFELLAPSSLRPVDMLLESMF VTPNTMASVSTVOLWISGLLALRVLISQSTED IVLSRIQGLSFSPYLISCTVINRLDGDSTSTILE BHSGRQUSN.PEETTSRPLQL.VGILLEDIVT KQLKVEMBEQQHTFYCQELGTLLMCLHIFKS GMFRRITAATATRLFRSDGCGGSFTLISDINLR ARSMITHPALVLLWCQILLLVNHTDVRWW ARVQCTYRKRISLSSTKLLSPQNAGGEEDSDILA ARLGMCNREIVRRGALLHCDVYCQNLHDSS HLYWINNHQDLISSSFKLLSPQNAGGEEDSDILA ARLGMCNREIVRRGALLHCDVYCQNLHDSS HLYWINNHQDLISSSFKLLSPQNAGGEEDSDILA ARLGMCNREIVRRGALLHCDVYCQNLHDSS HLYWINNHQDLISSSFKLLSPQNAGGEEDSDILA ARLGMCNREIVRRGALLHCDVYCQNLHDSS HLYWINNHQDLISSSPPLDSTARAMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQLSPS PPVSSIPLLBGDGHVSLETVSPDKDWYVHLVX SQCWTRSDSALLGGALNNIPRADMAFM MISERILSLAPCLS.GGMSBISGGQKSALFEA AREVTLARVSGTVOQLAVHHVPQFLPAPP AAYWSLANDLGGAALTVRVSTURDMARNSVYA MISERNISLLAPCLS.GGMSBISGGYSALFEA AREVTLARVSGTVOQLAVHHVPQFLPAPP AAYWSLANDLGGAALTVRVSTURDMARNSVYA SWSLADBGGAALTVRVSTURDFRIKX SWHLHBQPILS.DLQGAGLGCCLA.QLPCIL WSVYSTEPTHACSLLYCHFURLEAVAVQFG EQGFFEEDISRTQINVLAVQAITSLVLSAMT VPVAGNAVSCLEGQPRINCHLARNSCVYA PVSALDFRANTYRUNGTRELLARV			1				1
LKLQERVLNNVVIHLLGDEDPRVRHVANADQSSVYL KILIMIETQPPSHESVSTITRIYRGYNLLPSTID VITMENNLSRVIAAVSHEITSTTRALTFGCCE ALCILSTAFPVCIWSLGWHGVPPLSASDESR KSCTVGMATMILTILSSAWFPLDLSAHQDAL LLAGNILAASAPKSLRSSWASEEBANPAATK QESVWPALGDRALVPMVRQLFSHLLKVINIC AHVLDDVAPGFAKAALPSLTNPSLSPIRK GKEKPPGGASVPLSPKKGSEASAASRQSTST GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSSYRGLYHLCSYNLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFTALDLALDA SLRNMVQAGGENDTSGWFDVLQKVSTGLADA SLRNMVQAGGENDTSGWFDVLQKVSTGLADA SLRNMVQAGGENTSGWFDVLQKVSTGLADA SLRNMVQAGGENTSGWFDVLQKVSTGLADA GREGESFAGLENLTGPTWALADA GREGESFAGLENLTGPTWALADA CHKENEDK WKRLSRQIADILPMLAKQQMHH DSHBALGVINTLFELLAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTEDE VISSRIGGLSSFSYLJSCTVINRLRDGDSTSTLE EHSEGGQKNLPEETFSRFLLQLVGILLLCHIFKS GMFRRITAAATRLTRSDGCGGSFYTLDSLNL RSMTTHALLVLLVQCULLLVMHTDVRWW APVQQTPKRISLSSTKLLSPQMSGEEEDSLA AKLGMCNEEVRRGALHLFCDVCQLHDSS HLTWLINNHUPOLISLSHEPPVQDETSAVIRIS AASGLFQAIGSRCENLSTPTMAKTIQCLEDI HLSQSGAVLTLYVDRLLCTPFRVLARMYDLL ACRRVEMLLAANLQSSMAQLFMEELNRIQGS PPVSSIPLJGDGCHVSLETVSPDKDWYHVLW SQCWTRSDSALLEGAELVNNETTAMQDSLSPS PPVSSIPLJGDGCHVSLETVSPDKDWYHVLW SQCWTRSDSALLEGAELVNNETTAMQDSLSPS PPVSSIPLJGDGCCCLALQLPG MYSKSLDDGGAALYQSLFTLARALAQY LVVSKLPSHLHLPEREKDIVKYVATLEAL SWHLHEDPLSJDLQGAGLCCCLALQLPG GULSPFRETUTRALLGYCYLVTQFLVWLG GESPFEEDTERVNIVLVAQAITSLVLSAMT VPVAGNAVSKCLBGQPRINCHLARNSCVPA TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY							SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
LKLQERVLNNVVIHLLGDEDPRVRHVANADQSSVYL KILIMIETQPPSHESVSTITRIYRGYNLLPSTID VITMENNLSRVIAAVSHEITSTTRALTFGCCE ALCILSTAFPVCIWSLGWHGVPPLSASDESR KSCTVGMATMILTILSSAWFPLDLSAHQDAL LLAGNILAASAPKSLRSSWASEEBANPAATK QESVWPALGDRALVPMVRQLFSHLLKVINIC AHVLDDVAPGFAKAALPSLTNPSLSPIRK GKEKPPGGASVPLSPKKGSEASAASRQSTST GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFYHLSSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSSYRGLYHLCSYNLKATHA NYKVTLDLQNSTEKFGGFLSSAASRQSTST GPVTSKSSSLGSFTALDLALDA SLRNMVQAGGENDTSGWFDVLQKVSTGLADA SLRNMVQAGGENDTSGWFDVLQKVSTGLADA SLRNMVQAGGENTSGWFDVLQKVSTGLADA SLRNMVQAGGENTSGWFDVLQKVSTGLADA GREGESFAGLENLTGPTWALADA GREGESFAGLENLTGPTWALADA CHKENEDK WKRLSRQIADILPMLAKQQMHH DSHBALGVINTLFELLAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTEDE VISSRIGGLSSFSYLJSCTVINRLRDGDSTSTLE EHSEGGQKNLPEETFSRFLLQLVGILLLCHIFKS GMFRRITAAATRLTRSDGCGGSFYTLDSLNL RSMTTHALLVLLVQCULLLVMHTDVRWW APVQQTPKRISLSSTKLLSPQMSGEEEDSLA AKLGMCNEEVRRGALHLFCDVCQLHDSS HLTWLINNHUPOLISLSHEPPVQDETSAVIRIS AASGLFQAIGSRCENLSTPTMAKTIQCLEDI HLSQSGAVLTLYVDRLLCTPFRVLARMYDLL ACRRVEMLLAANLQSSMAQLFMEELNRIQGS PPVSSIPLJGDGCHVSLETVSPDKDWYHVLW SQCWTRSDSALLEGAELVNNETTAMQDSLSPS PPVSSIPLJGDGCHVSLETVSPDKDWYHVLW SQCWTRSDSALLEGAELVNNETTAMQDSLSPS PPVSSIPLJGDGCCCLALQLPG MYSKSLDDGGAALYQSLFTLARALAQY LVVSKLPSHLHLPEREKDIVKYVATLEAL SWHLHEDPLSJDLQGAGLCCCLALQLPG GULSPFRETUTRALLGYCYLVTQFLVWLG GESPFEEDTERVNIVLVAQAITSLVLSAMT VPVAGNAVSKCLBGQPRINCHLARNSCVPA TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY TAACEGREEGEDFRONTONTRY							LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
IRLUPKLIFYKCDQGQADPVYAARQOGSVYLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE LICLUSTAPPVCIPSI.G WHGGVYPLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAPPVCIPS.G WHGGVYPLSASDESR KSCTVGMATMILTILSSAWPPLDLSAHQDAL ILAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVPMVPQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTWPPSLSPIRWA (REKERGGRASV)SPKKSGSASASRGSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHAA NYKVTLDLQNSTEKFGGFLRSALVSQULE ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFPGLSSNFSKSGGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRINNYQAGEDDISTSWFDVLQKVSTQLKT NLTSVTRNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVDLLALQVLQTLAVYCLL DSDQVFIGFVLKQFEYTEVGGPRESEAHPMILV DSDQVFIGFVLKQFEYTEVGGPRESEAHPMILV DSDQVFIGFVLKQFEYTEVGGPRESEAHPMILV DSDQVFIGFVLKQFEYTEVGGPRESEAHPMILV CCHKENEDKWKLSRQLADDILPMLASKOPH LYSVERYHSKQHGPKIQLCGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LTCQKEVVSMLLKLIQYHQVLEHIVLOQ CIKENEDKWKLSRQLADDILPMLASKOPH VTPNTMASVSTVOLWISGLIALRVLISGSTSTLE ULSUGLESFSPVLISCTVINRLRDGSTSTLE UNLSRIGGLSFSPVLISCTVINRLRDGSTSTLE EHSEGKQLENLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLLIHEKS GMFRRITAATALRFSDGCGGSFTLDSLNILR ARSMITHFALVLLWCQULLLVNHTDVRWW ARVQQTTKRHISLSSTKLLSPQNAGSEEDSDLA ARLGGKONEIVRRGALLICCDYVCQNLHDSS HLTWLINNHQDLISSSFELLJSCHPONGSEEDSDLA ARLGGKONEIVRRGALLICCDYVCQNLHDSS HLTWLINNHQDLISSSFELLJSCHPONGSEEDSDLA ARLGGKONEIVRRGALLICCDYVCQNLHDSS HLTWLINNHQDLISSSPEPVOSPILDSLNIRN AASGLFIQAJQSRCENLSTPMLKKTLQCLEG HLSQSGAVITLYVDRLLCTFFRYLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELINIQES PVSSIPLJBGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLGGALVNRIPAGDMAFM MISERIUSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVOQLPAVHHVPQPLPAPA AYWSLANDLEGDAALYQSLPTLARALVQDG UNVSYLPSHLHLIPPEREDDIVKVAVICAL SWHLHEQPILSJLDQGAVLAVQPG PVSLIPPEREDDIVKVAVICAL SWHLHEQPILSJLDQGAVLAVQPG AYWSLANDLEGDAALYQSLPTLARRALVQDG WSVYSCHPSHLALLPREEDINKVAVQFTEVHVKIR GESPPEEDTERTOINVLAVQAITSLVLVSMT TAACCMYARMYSCLQOPRINCHLAALAVQPG PVSLSPATTGALISCHELLLQNPFLYWKIR GESPPEEDTERTOINVLAVQAITSLVLYGHVM GESPPEEDTERTOINVLAVQAITSLVLVSMT VPVAGNAVSCLEQQPRINCHEKLALDTRFGKK LSIIRGIVEGEIQAMVSKRENNATHHHLYQAWD PVSLSPATTGALSHEELLLQUPFEELGSMS	1						
KLLMHETQPPSHPSVSTITRIVRGYNLLPSTID VTMENNLSRVIAAVSHELITSALFGCCE ALCLISTAFPVCIWSLGWIGGYPPLASADESR KSCTVGMATMILTLLSSAWPLDLSAHDDAL ILAGNILAASAPKSLRSSWASEEEANPAATK QEEVWPALGDRALVPWPEQLFSHLKVNIC AHVLDDVAPEGAIKAALPSITNPPSLSPIRK GKEKEPGOASVPLSPKKGSEASAASKQSDTS GPVTTSKSSLGSYPHLPSVLKHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCTSREPMMATVC VQQLKTLFGTNLASQPDGLSSNPSKSOGRA QRLGSSSVPGLTHYCFMAPYTHFTQALDA SLRNMYQAEQENDTSGWPDVLQKVSTOLKT NLTSYTKNRADKANENHHIRFELFLVKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVEIGFVLKQFEVEWCGPESALIPNIFF FLVLLSVERYPSKQIIGGKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETTQREVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDEKWKRLSRQIADIILPMLAKQQMHI DSHEALGYLNTLFEILAFSSLAVBLSGTED IVLSRQELSFSPYLISCTVNRLRDGDSTSTLE EHSEGKQKKNLPETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELTULLUHINTDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTFKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLVHITDYRWW AEVQQTPKRHSLSSTKLLLSPOMSGEEDSDLA AKLGMCNREIVRRGAALLFCVONLHDSE HLTWLLTWHIQDLISLSHEPPVQDFISAVHRNS AASGFIQAUGSCENL STFTMKLTLQCLIS GHLSQGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLAANALGSSMAQLPMELNRIQEY LQSSGAVGRISDSLALBGABLLVNRPALSTMQDSLSP PVSSHLDGDGHVSLETVSPDKDWYHLVK SQCWTRSDSALLBGABLLVNRPAEDMAMFM MNSEPNLSLLAPCLSLGMSEISGGKSALIFFA AREVTLARVSGTVQLLAVHVPQFELPAEP AAYWSKLNDLFGDAALYQSLPTLARALQQY LVVVSKLPSHLHLPPEKEKDVFVATLEAL SWHLHBGOFISLDLQAGLDCCLALQPGL WSVVSSTEFVTHACSLIVCVGHLEAVAVQPG EQLLSPERGTTMTHALSGEEEEVDPNTONEYN TAACEMVAEMVESLOSALGHRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLWWLG WSRFROOFGAFPETFVETLQKEVFKEFTRR INTIGWTSRTQFETFWATLLGALFATRAVQPL GESSPFEEDTERTONLAVAJOTISLUSAMT VFVAGNPAVSCLEQQFRNRYKALADTRFGKK LSHIGHIVGOEGQAMYSKRENIATHHLLYQAWD PVPSLSSATTGGLISHEKLLUGNPERELGSMS							
VTMENNI.SRVIAAVSHELITSTIRALTFGCCE ALCLISTAFPCIWSLGWIGGYPPLSASDESR KSCTVGMATMILTILSSAWFPLDLSAHQDAL ILAGNILAASAPKSLRSSWASEEANPAATK QEEVWPALGDRALVFMYEQLFSHLKKVINIC AHVLDDVAPGAKAALPSLTNPPSLSPIRK GKEKEPGEDASVPLSRKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTDLQNISTEKFGGFIRSALDVLSQILEL ATLODIGKCVBEILGYLKSCFSREPMAATVC VQQLKTLFGTILASQPGDESSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALDA SLRMMYQAEQENDTSGWFDVLKYSTOLKT NLTSYTKNRADKNAHNHIRLFEPLVIKALKO YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFYLKQFEYEVGQFRESEAILPNIFF FLVLLSVERYFSKQIIGGFKIIQLCDGIMASGR KAVTHAPALQYPHOLIPLVIRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CIKENDEW KRRLSRQIADILARQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGLALRAVLISQSTED IVLSRQIELSFSPYLISCTVINRLRDGDSTSTLE EISGGKQIKNLPETFSRFLICGILEDIVT KQLKVEMSEQQITFYCQELGTLLMCLDIEKS GMFRITAAATRIFRSDGCGGSFYTLDSILNLR ARSMITHPALVLLWCQILLLVNITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSLA AKLGMCNREIVRGAALLLVNITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSLA AKLGMCNREIVRGAALLLCVNITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSLA AKLGMCNREIVRGAALLCVNITDYRGVW AEVQGTPKRHSLSSTKLLSPQMSGEEEDSLA AKLGMCNREIVRGAALLCVNITDYRGWW AEVQGTPKRHSLSSTKLLSPQMSGEEEDSLA AKLGMCNREIVRGAALLCVNITDYRGWW AEVQGTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRGAALLCSNMADPELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGIVSLETVSPRELNRIGEY LQSSGLAQRHQRLYSLLDRFRLSKRUYALIGHRNSGVPA AYWSKLNDLFGGAKALGHKNSGVPA FLTPLLRMISLARLPLNKALDTRGKK SWHLIHEQIPLSLDLQAGLDCCCLALQCROPNKYH NTLGWTSRTQFEETWATLLGVLVTOPLVMLG GESPPEEDTERTONNLAVQATISLULSAMT VYVAGNPAVSCLEQQPRINKPLKALDTRGKK LSIIRGIVEGEIQAMYSKENIATHHLYQAWD PYPSLSSATTGGLEGGMAYSKENIATHHLYQAWD PYPSLSATTGGLISHEKLLLQIPRORLGAWD	1						
ALCLISTAFPYCIWSLGWIEGOPPLASABDESR KSCTVGMATMILTLISSAW PREDLISAHODAL ILAGAILAASAPKSLRSSW ASEEEANPAATK QEEVWPALGDRALVPMVBQLFSHLLKVINIC AHVLDDVAPGAIKAALPSITNPPSLSPIRKK GKEKEPGQASVPLSPKKGSEASAASKQSDTS GPVTTSKSSLGSFYHLPSYLKHDVLKATHA NYKVTLLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGK VEBELIGYLKSCFSREPMMATVC VQOLKTLFGTINLASQFDGLSSNPSSSQGRA QRLGSSSVPROLYHYCFMAPHTQALDA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYTRIRADKNAINHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVPIGFVLKQFEVIEVGQFRESEAIPNIFF FLVLLSVERYTBSQIIGGPKIJQLCOGIMASGR KAYTHAIPALQPIVIEIDLFVLRGTNKADAGKE LEFIQKEVVSMILRILQYPIQUEMPFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMH DSHEALGYLNTJFEILAPSSLAFVOMLLRSWF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRQIBLSSPYPLISCTVINRLEDGDSTSTLE EHISGKQIKNLFEETFSRELQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIEKS GMFRRTAAATRLFSBCGGCGTLJQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIEKS GMFRRTAAATRLFSBCGGCGFKIJLVGUHDEN ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQOTPKRHSLSSTKLLSPGMSGEEDSDLA ALLGMCNREIVRRGALLIFCDYVCONLEDSE HLTWLNYHIDDUISLSHEPPUQPISAVPRW ARSMITHPALVLLWCQILLLVNHTDYRWW ARSWGTRFRSBALLGCAGCGSALFEA ARSGAVTLAVSGTRORUSELSTSTLSPGMSGEEDSDLA ALLGMCNREIVRRGALLIFCDYVCONLEDSE HLTWLNYHIDDUISLSHEPPUQPISAVPRW ASSAGLIQAIQRHQRLYSLLDRFRLSTMODSLSPS PPVSSHPLOGGDGIVSLETVSPBLDWYVHLVX SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSENISLLAPCLS GMSEGEDSDLA ARSGAVITLAVNGTURGLARLIPCDYVCONLEDSE HLTWLNYHIDDUISLSHEPPUQPISAVPRMS ARSGYTLARVSGTVQQLPAVPHVPQPELPAEP AAYWSKLNDLFGDAALVGSLVKFVVATLEAL SWHLIBEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIVCVHFLEANAVQPG EQULSPBRRTNTRKAISEEEEEVDPNTOMKYI TAACEMVAEMVSELQOPANCHYKALDTRGKK LSIBGIVEGEIGAMNSKENIATHHLYQAWD PYPSLSARTGEGEIGRAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD PYPSLSARTGEGORAMNSKENIATHHLYQAWD							
KSCTYGMATMILTILLSSAWFPLDLSAHQDAL ILAGNILAASAPSCRISSWASPALDRATK QEEVWPALGDRALVPWYEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKBPGEQASVPLSPKKOSEASAASRQSDTS GPVTTSKSSSLGSFYHLBSVLKLHDVLKATHA NYXVTJLDLQNSTERFGGFIRSALDVLSQILEL ATLQDIGKCVBEILGYLKSCFSREPMAATVC VQLLKTLFGTILASQFOELGSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRINMYQAECBHDTSGWFDVLGKVSTQLKT NLTSYTKNRADKNAINHIRLFEPLVIKALKQ YTTTTCVQLQKVJLLLAQLVQLRVMYCLL DSDQVFIGFVLKQFEYBEVGQFRESEAIIPNIFF FLVLLSYBRYTSKQIIGIPKIQLCOGIMASGR KAVTHAPIALQPYIVIDLEVURGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CIKKENDEWKRKLSRQIADILAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTYQLWISGILAKQVGNFILSTLE HESGKQIKNLPETFSRFLQLVGILLEDIVT KQLKVPMSPQQHTFYCQELGTLLMCLIHEKS GMFRRTAAATRLFRSDGGGSFYTLDSILNR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSQMSGCEEDSDLA AKLGMCNBEIVRGAGLLSTLLSPVQNLHDSE HLTWLIVNHIQDLISLSHEPYQDFISAVRNS AASGLFQAIQSGCENLSTTHKKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVBELLAANLQSSMAQLYRNS AASGLFQAIQSGCNLSTTHKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVBELLAANLQSSMAQLYRNS AASGLFQAIQSGCNLSTTHKKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVBELLAANLQSSMAQLPBLINDIGEY LQSSGLAQRHQRLYSLLDFFRI STMODSLSPS PPVSSHPLOGGGPHSLETVSPEDKDWYWHLVK SQCWTRSDSALLEGABL VNRIPAEDMNAFM MNSEPNISLLAPCLSL GMSEISGGVSALFEA AREVTLARVSGTVQQLPAVHIVPQPELPAEP AAYWSKLNDLFGDAALYQSLPGLARALQY LVVXKLPSHLHLPPEKKEDVEVAVTLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVVSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALQPGL WSVVSSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALQPGL WSVVSSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALQPGL WSVSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALAQPG LWSVSSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALAQPG WSVRSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALAQPGL WSVSTEFVTHACSLUCVEHILBAVAVQFG EQLLSPERRITTHKASTGLOGCLALAQPGLEXCHYKLFIYR NTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTONLAVAQTISLUCXAMT VYVAGNPAVSCLEQQPRINKLTKALDTRGKK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PYPSLSATTGAGLISHEKLLLQHANDSM							VTMENNLSRVIAAVSHELITSTTRALTFGCCE
KSCTYGMATMILTILLSSAWFPLDLSAHIODAL ILAGNILAASAPSLRISWASPALPAATK QEEVWPALGDRALVPWYEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEQASVPLSPKKOSSASAASRQSDTS GPVTTSKSSSLGSFYHLESVLKLHDVLKATHA NYXVTJLDLONSTEEFGGFIRSALDVLSQILEL ATLODIGKCVBEILGYLKSCFSREPMMATVC VQLLKTLFGTILASQPGDLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQALADA SLRINMYQAREGEDTSGWFDVLQKVSTQLKT NLTSYTKNRADKNAIHHIRLFEPLVIRALDQ YTTTTCVQLQKVJLLLAQVQLVAVYCLL DSDQVFIGFVLKQFEYBEVGGFRESAHIPNIFF FLVLLSYBRYBSKQHIGHKIQLCOGIMASGR KAVTHAPIALQYHVILDEVURGTNKADAGKE LETQKBVVVSMLLRLIQYHQVLEMFILVLQQ CIKKENDEWKRKJSRQJADILJPMLAKQQMHI DSHEALGYLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTYQLWISGLALIKLISQSTED IVLSRIQBLSFSPYLISCTVINRLRDGDSTSTLE BHSGKQIKNLPETFSRFLQLVGILLEDIVT KQLKVPMSPQQHTFYCQELGTLLMCLBIBKS GMFRRTAAATRLFSDGGGSFYTLDSTLNR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNBIDYRGAALLSCSMAQLYNRNS AASGLFQAJQSGCENLSTITMIKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVBELLAANLQSSMAQLFBLINKIKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACREVBELLAANLQSSMAQLPBLINKICFY LQSSGLAQRHQRLYSLLDFRLSTMGDSLSPS PPVSSHPLDGGBIVSLETVSPBLDWVYHLVK SQCWTRSDSALLBGABLVRIPAEDMNAFM MNSEPNLSLAPCLSLGMSEISGGVSALFEA AREVTLARVSGTVQQLPAVHFUPPELPAEP AAYWSKLNDLFGDAALYQSLPGLARGL USWLSTEFYTHACSLIVCYVATLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVVSSTEFYTHACSLIVCYVATLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVVSSTEFYTHACSLIVCYVATLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVSSTEFYTHACSLIVCYVATLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVSSTEFYTHACSLIVCYVATLEAL SWHLIHBQIPLSLDLQAGLDCCCLALQPGL WSVSSTEFYTHACSLIVCYWIFLEAVAVQF GQLJSPBRRTNTPKAISEEGEEVDPNTONINKYI TAACCENVAERMVSSLOQALLGHRNSGVPA FLTPLERNIISLARLPLYNSYTRVPPLVWILG WSVSSTEFYTHACSLIVCYWIFLEAVAVQFG GQLJSPBRRTNTPKAISEEGEEVDPNTONINKYI TAACCENVAERMVSSLOQALALGHRNSGVPA FLTPLERNIISLARLPLYNSYTRVPPLVWILG WSVSSTEFYTHACSLIVCYWILGAMTYONLAVAGITSLULSAMT VPVAGNPAVSCLGQPRINKPLKALDTRGKK LSIIRGIVEGEIQAMYSKRENIATHHLYQAWD PYPSLSSATTGGLISHEKLLLQHRNSGVPA							ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
ILAGNILAASAPKSLRSSWASEEANPÄATK QEESWPALGDRAL/PMPGLSHILK/VINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGCASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKHHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATILQDIGKCVERILGYLKSCSSREPMMATVC VQULKTLFGTNLASQFDGLSSNPSKSQGRA QRIGSSSVPRGLYHCYFMYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKYSTQLKT NLTSVTKNRADKNAHTHRILFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYTEVGQFRESEAIPNIFF FLVLLSYERYHSKQUIGPKIQLCDGIMASGR KAVTHAPALQPIVIDLFVLRGTNKADACKE LETQKEVVYSMLLRLIQVHQVLEMFILVLQQ CHKENEDKWRLSSQIADIILPALAQQMHI DSHEALGVLNTLFELLAPSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETTSRFLQLVGILLEDINTY KQLKVEMSEQQHTFYCQELGTLLMCLIIIFKS GMFRRTAAATBLFRSDCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSGGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSGWGSEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIODLISLSHEPPYQDFISAVYRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEEINIQEY LQSSGLAQRHQRLYSLLDRFRLSTMDSLSPS PPVSSHPLDGDGHYSLETVARMVDLL ACRRVEMILAANLQSSMAQLPMEEINIQEY LQSSGLAQRHQRLYSLLDRFRLSTMDSLSPS PPVSSHPLDGDGHYSLETVARMVDLS SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPILSLAPCLSLGMSGGGKSALFEA AREVTLARVSGTVQQLPAVHHYPQPELPAEP AAYWSKLDDFGDGHYSLETVIAALAQP ALYWSKLDDLGGDGHYSLETVIAALAQP LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHBGQIPLSLIPTLARALQP AAYWSKLDDLFGDAALVSTLAPALAQP ALYSSLAPLVSTLAPLAQP ALYSSLAPLYNNSTRYPQINGNEYI TAACEMVAEMSCLGQFRKYRLKALDLPFGRK NTLACKHYNSTRYPLWKLI WVSNSTEFVTHACSLIFVCHFLEAVAVQPG EQLLSPERTTTTRAISEEEEEVDPTNQNRYI TAACEMVAEMPSCLGQFRKPLKALDLPFGRK NTLGWTSRTVTRAISECEEEUPTNQNRYI TAACEMVAEMPSCLGQFRKPLKALDLFGGK STURTINGTSRTVTRAISECEEEUPTNQNRYI NTLGWTSRTQFEETWATLLQVLVTQLVME QEESPPEEDTERTONVLAVQLAFGEKSKEFIY NTLGWTSRTQFEETWATLLGVLVTQLVME QEESPPEEDTERTONVLAVQLAFGEKSMS]						
QEEVWPALGDRALVPMYEQLESHLILKUNIC AHVLDWAPGPAKAAPISTNPPSISPIRRK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSTYHLPSYLKHUDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCFSERPMATVC VQQLLKTLGTNLASQFGGLSSNPSKSQGRA QRLGSSSVRPGLYHVCFMAPYTHFTQALADA SLRNMYQAGEONDTSGWDVLQKVSTOLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLENYCLL DSDQVFIGFVLKQFEYEVGGYRESEAHPNIFF FLVLLSYERYHSKQIGIPKIQLCDGIMASGR KAVTHAIPALOPIVHDLFVGTNKADAGKE LETQKEVVSMLLRLIQVHQVLEWFILVLQQ CHKENEDKWKRLSRQIDHENIQLCDGIMASGR KAVTHAIPALOPIVHDLFVGTNKADAGKE LETQKEVVSMLLRLIQVHQVLEWFILVLQQ CHKENEDKWKRLSRQIDHILDALAQQMH DSHBALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVILSQSTED IVLSRQELSFSPYLISCTVINRLRDGDSTSTLE EHSGGQIKTYCQCLGTLLMCLIHIEKS GMFRRITAAATBLFRSDGCGGSFYTLDSLNLR ARSMITTIPALVLILWQLILLDIVTY KQLKVEMSEQQHTFYCQCLGTLLMCLIHIEKS GMFRRITAAATBLFRSDGCGGSFYTLDSLNLR ARSMITTHAIVLVINHIDTYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCKREIVRRGALLLFCDYVCQNLHDSE HLTWLLVNHIQDLISLSHEPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACREVEMILAANLQSSMAQLPMEEINRIQEY LQSSGLAQRIQRLYSLLDFFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMAFM MNSEPHISLLAPCLSLGMSEISGGQKSALFDA AREVTLARVSGTVQQLPAVHHVFQELPAEP AAFWSKLNDLFGDAALYGSLPTLARALAQY LVVVSKLPSHLHILPEKEKDIVKFVAATLEAL SWH-IH-BQGPLSLJDLQGALVSLPFLARALAQY LVVVSKLPSHLHILPEKEKDIVKFVAATLEAL SWH-IH-BQFISLJDLQGALVSLPFLARALAQY LVVVSKLPSHLHILPEKEKDIVKFVAATLEAL SWH-IH-BQFISLJDLQGALVSLPFLARALAQY LVVVSKLPSHLHILPEKEKDIVKFVVATLEAL SWH-IH-BQFISLJDLQGALVSLPFLARALAQY LVVVSKLPSHLHILPEKEKDIVKFVVATLEAL SWH-IH-BQFISLSLDLQGALVSLPFLYWNICG WSFVPGGGDFTAFFEIFFFLQEKEFIYR NTLGWTSRTQFEETWATLQLV-VTPLWME QEESPFEOTERTONIVLAVQALTISLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAWSKRENIATHHLYQAWD PVPSLSRATIGAUSSERGEGNSEMS	1						
AHVLIDDVAPGPAIKAALPSLTNIPSLSPIRIK GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVSQILEL ATIQDIGKCVEBLGYLKSCFSREPMMATVC VQULKTLFGTNLASQFFGLSSNPSKSQGRA QRUGSSSVPRGLYHCYFMYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHIRLFPLVIKALKQ YTTTTCVQLGVQLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEWGGPRESEAIPNIFF FLVLLSYERYHSKQIGIPSHQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETQKEVVYSMLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMH DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTNTMASSVTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGRQIKNIPETTESRTLUVGILLGUFLLSVILSL GMRRRITAAATRLFSDGCGGSFYTLDSLINR ARSMITTIPALVLLWCQLLVVGILLEDIYK GUKVEMSEQQHTFYCQELGTLLMCLIIIFKS GMRRRITAAATRLFSDGCGGSFYTLDSLINR ARSMITHPALVLLWCQLLLVHNTDYRWW AEVQQTPKRHISLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRGALLIFCDYVCQNLLDSS HLTWLIVNHIODLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAGRRQLTSLLDFRALBRINGEY PVSSHPLDGGGHVSLETVSPDKDWYHLVX SQCWTRSDSALLEGAELVNRPAEDMMAFM MNSEFNLSLLAPCLS LGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHYFQELPAEP AAYWSKLNDLFGDAALYSQLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPTSLDLGQGLAVGHERNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSVVSSTEFVTHACSLLYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEEVDPPNTONPKYI TAACEMVAEMVESLQSVLAIGHKRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSVPSGGGTAFPEIFPVLSVLAGKHRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSVPSGGGTAFPEIFPVLSVLAGKHRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSVPSGGGTAFPEIFPVLSVLAGKHRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSVPSGGGTAFPEIFPVEFLQEKEVKFEFIYR INTLGWTSRTOPEETWATLLGVLYTOPLVMG QEESPPEEDTFTAOLVAQAITSLASAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFCRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSRATTAGLSLLLVQNPFRELGSMS]						
GKEKEPGEOASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYH.PSYLHDVLKATHA NYKVILDLQNSTEKFGGFI.RSALDVLSQILEL ATIQDIGKCVEELIGYIKSCFSEPPMATVC VQQLLKTLFGTNLASOPGI.SSNPSKSQGRA QRLGSSSVRPGLYHVCPMAPYTHFTQALADA SIRNMYQAGEONDTSGWPULGKVSTQLKT NLTSVTKNRADKNAJHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQEYTEVGGRESEAIIPNIFF FFULLSYERYHSKQIGIPKIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEHPILVLQQ CHENEDSWKRLSRQIDJILPMLAKQQMH DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAIRVLISQSTED IVLSRQELSFSPYLISCTVINRLRDGDSTSTLE EHSGGQIKN.PEETFSRFLLQLVGILEDIVT KQLKVEMSEQQHTFYCQFLGTLLMCLIIIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTIPALVLLWCQILLLVHTIDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRGGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSFEPYQPFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLTVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQLYSLLTVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRLYSLLDFFRSTLSMODSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYYHLVX SQCWTRSDSALLEGAELVNRIPAEDMMAFM MNSEFNLSLLAPCLS GREISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVPGELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLBLEAGELVNRIPAEDMMAFM MNSEFNLSLLAPCLS GREISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFGLPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLBLEPEKERIDVGFVAATLEAL SWHLBFEQIPLS LDLQAGLGCCLALQLPEL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTTITKRAISEEGEEVDPTNQNRYI TAACEMVAEMYESLGSVLALGHKRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWLG WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTTITKRAISEEGEEVDPTNQNRYI TAACEMVAEMYESLGSVLALGHKRNSGYPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWLG GESPPEEDTERTONIVLAVQAITSLVLSAMT VPVAGPPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSRATTGALISHERLLLQNPFERELGSMS					İ		•
GPVTTSKSSLGSFYHLPSYLKLHDVILKATHAA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATIQDIGKCVEELGYLKSCSSREPMMATVC VQQLLKTLFGTINLASQFDGLSSNSKSQGRA QRLGSSSVRGU,HYLQFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLKQVSTQLKT NLTSVTKNRADKNAHBIRJEFPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEVEVGGFRESAHIPNIFF FLVLLSYSEYHSKQIGFKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTINKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQJADHLPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRIQELSFSYVLISCTVINRLRDGDSTSTLE EHSEGKQUKNLPEETFSFFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLLVGMILEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLLVGMILEDIVT KQLKVEMSEQQHTFVCQDELGTLMCLHIFKS GMFRITAAATRLFSSFLLLSQNMSGEEDSDLA AKLGMCNEIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPVQDFISAVHRNS AASGLFQAQNSRCENS.STFMLKKTLQLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMQLLPMELNRIQEY LQSSGLAQRIQRIVSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGFIVSLETVSPDKDWYVHLVK SQCAVTRSDSALLEGARVNIRNAEDMNAFM MISEPNLSLLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQFELPAEP AAYWSKLNDLEGDAALYQSLFTLARALAQY LVVVSKLPSHLHLPFEKEDIVKFVATLEAL SWHLIHBGPISLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTFKAISEEEEVDPNTQNPKYI TAACEMVAEMVESILGSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTR VPPLVWKLG WSPKPGGDFGTAFFEPFLQEKEVKKEFIVK NITLGWTSRTOFFETWATLLLGVLTQPLVME QEESPPEEDTERTQNIVLAVQATISL ULSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMYSKRENIATTHLYQAWD PVPSLSPATTGALISHEKLLLQNNFERELSGMS	1						AHVLDDVAPGPAJKAALPSLTNPPSLSPIRRK
GPVTTSKSSLGSFYHLPSYLKLHDVILKATHAA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATIQDIGKCVEELGYLKSCSSREPMMATVC VQQLLKTLFGTINLASQFDGLSSNSKSQGRA QRLGSSSVRGU,HYLQFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLKQVSTQLKT NLTSVTKNRADKNAHBIRJEFPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEVEVGGFRESAHIPNIFF FLVLLSYSEYHSKQIGFKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTINKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQJADHLPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRIQELSFSYVLISCTVINRLRDGDSTSTLE EHSEGKQUKNLPEETFSFFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLLVGMILEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFSSFLLLVGMILEDIVT KQLKVEMSEQQHTFVCQDELGTLMCLHIFKS GMFRITAAATRLFSSFLLLSQNMSGEEDSDLA AKLGMCNEIVRRGALLFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPVQDFISAVHRNS AASGLFQAQNSRCENS.STFMLKKTLQLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMQLLPMELNRIQEY LQSSGLAQRIQRIVSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGFIVSLETVSPDKDWYVHLVK SQCAVTRSDSALLEGARVNIRNAEDMNAFM MISEPNLSLLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQFELPAEP AAYWSKLNDLEGDAALYQSLFTLARALAQY LVVVSKLPSHLHLPFEKEDIVKFVATLEAL SWHLIHBGPISLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTFKAISEEEEVDPNTQNPKYI TAACEMVAEMVESILGSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTR VPPLVWKLG WSPKPGGDFGTAFFEPFLQEKEVKKEFIVK NITLGWTSRTOFFETWATLLLGVLTQPLVME QEESPPEEDTERTQNIVLAVQATISL ULSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMYSKRENIATTHLYQAWD PVPSLSPATTGALISHEKLLLQNNFERELSGMS	1.	1					GKEKEPGEQASVPLSPKKGSEASAASROSDTS
NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTINLASQFDGLSSNPSKSGGRA QRLGSSSVRFGLVHYCFMAPYTHFTQALADA SLRNMVQAEQENDTSGPVJLQKVSTQLKT NLTSVTKNRADKNAIHHHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEVEVGGFRESEAIPNIFF FLVLLSYERYHSKQIIGFKIQLCDGIMASGR KAVTHAIPALQPVIPUDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAFSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRIQGLSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEFTSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTILLMCLIIIFKS GMFRITAAATRLFSDGCGGSFYTLDSLNLR ARSMITTHALVLLWCQILLLVNHTDYRWW AEVQQTYRKHSLSTSKLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTFTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHGRLYSLLDRFRLSTMODSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGGALLVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGQKSALFEA AREVITLARVSGTVQQLJAVHHVFQPELPAEP AAYWSKLNDLFGDAALLYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVFVVATLEAL SWHLHBEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG GQLLSPERTNTPKALSEEEDEDDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FILTPLLRNIIISLARLPLVNSYTRYPLVWKLG WSPKPGGDFGTAPFEPVEFTQEKEVKEFTVR NTLGWTSRTQFEETWATLLGLVTQPLVME QESSPPEEDTERTQNVLAVQAJTSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMYSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPGELGSMS	1						GPVTTSKSSSLGSFYHT PSYLKLHDVLKATHA
ATLODIGKČVEELIGYLKSCFSREPMMÄTVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHIFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHHIRLEFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEVGQFRESEALPINIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVIIDLEVLRGTNKADAGKE LETQKEVVSMLLRLQYHQVLEMFILVLQQ CHKENEDKWKRLSRQADILIPMLAKQQMHI DSHBALGVLNTLFEILAPSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED INLSRQELSFSYLISCTVINRLTDGDSTSTILE EHSSGKQIKNLPESTFSFLLQLVGIILLEDIVT KQLKVEMSEQQHTFVCQELGTILMCLIHIFKS GMFRRITAAATRLFSSEGCGGSFYLLDSLNLR ARSMITTHPALVLLWCQILLLWHTDYRWW AEVQQTPKRHSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGGALLICDVYCQNLHDSE HLTWLIVNHQDLISLSHEPPVODFISAVHRNS AASGLFQAIQSRCENLSTFITMLKKTLQCLGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDII. ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRRQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGALVNRPAEDMNAFM MISSFNLSLAPCLSLGMSEISGGGKSALFEA AREVITLARVSGTVQQLPAVHIVFQPELPAEP AAYWSKLNDLFDGAALQSLTTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVATLEAL SWHLHBEQPISLDLAQDCCCLALQLPGL WSVVSTEFVTHACSLIYCVHFILEAVAVQPG GQLSPSFERTNTPKAISEEEEDDFDFTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLIPLLRNIISLAARLP.VNSYTRAPLOWNRKJ WSPKPGGDFGTAFFEPVEFLQEKEVKKETVK NITLGWTSRTOFFETWATLLLGVLTOPLVME QEESPFEEDTERTONVLAVQAITSLUTSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEQAMVSKRENIATHLYQAWU PVPSLSPATTGALISHEKLLLQNPERLGSMS	1						•
VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRINMVQAEQENDTSGWFDVLQKVSTQLKT NILTSVTKNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLLLAQLVQLRVNYCLL DSDQVPIGFVLKQFEYEVGGFRESEAHPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTTKKADAGKE LETQKEVVVSMLLRLQYYHQVLEMFILVLQQ CHKENEDKWRKLSRQIADIILPMLAKQOMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLINHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLSSMAQLPMEELINRIGEY LQSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGDHVSLTVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MINSEFNLSLLAPCLSLGMSEISGGKSALFEA AREVILARVSGTVQLFVHHVPQFELPAEP AAYWSLINDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHBQIPLSLDQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCWFILEAVAVQPG EQLSPPERTNTTKAISEEEEEVDPTYONPKYI TAACEMVAEMYESLQSVLALGHKRNSGYPA FILPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKEFIVX INTLGWTSRTQFEETWATLLGVLVTQPLVME QESPPEEDTERTQINAVAVQAITSLUSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVEQEIQAMVSKREINTHHLYQAWD PVPSLSPATTGALISHEKLLLQNPTERLGSMS							,
ORLGSSSVRPGLYHYCFMAPYTHIFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVPIGFVLKQFEVVGGFRESEAIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLHGVYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSBGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTTYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPOMSGEEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTIMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARWDDL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPBKDWYVHLVK SQCWTRSDSALLEGABLVNRIPAEDMNAFM MISETPILSLAPGLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYGSLFTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKRVVATLEAL SWHLHBGJPLSLDLQAGLDCCCLALQLPGI WSVVSSTEPTHACSLIYCVHFILEAVAVQPG EQLLSPERRTINTFKAISEEEEEVDPNTQNFKYI TAACEMVAEMYESLQSVLALGHKRNSGVPA FLIPLLERNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR NTLGWTSRTTQFEETWATLLGVLVTQPLVME QESPPPEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRINFLKALDTRFGKK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALLISHEKLLLQINPERELGSMS	1 1						
SLRIMMYQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIHNHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQEFYEVGGFRESEAIIPNIFF FLVLLSYERYHSKQIGIPKHIQLCOGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWSGILAIR.VLSQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHFTYCQELGTLLMCLUHIFKS GMFRRTAAATRLFRSDGGGSFYTLDSLNLR ARSMITTIPALVLLWCQILLLVHHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISHEDPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGGHVSLTVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSLAPCLSLGMSEIGGGKSALFEA AREVTLARVSGTVQQLPAVHHVPQPELPAEP AAYWSKLNDLFGDAALYGSLPTLARALAQY LVVSKLPSHLHIPERKEKDIVKEVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEANAVQPG EQLLSPERRTNITRAISEEEEEVDPTNQPKYI TAACEMVAEMYESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKEFIVK NTLGWTSRTQFEETWATLLGGLVTQPLVME QESPPPEDTERTQINAVAQAITSLUSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPTERCLGSMS	i i						
NILTSVTKÜRADKNAHNHIRLFEPLVIKÄLKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFULAQFEYTEVGGYRESSAIIPNIFF FLVILSYBEYHSKQIIGFBYIQLCDGIMASGR KAVTHAPALQPVHDLFVLRGTIKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQLADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRYEMILAANLQSSMAQLPMEEINRIGEY LQSSGLAQRHQRLYSTLDRFFRVLARMVDIL ACRRYEMLLAANLQSSMAQLPMEEINRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMAFM MNSEFNLSLLAFLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYGSLPTLARLAQY LVVVSKLPSHLHPERKEDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLLYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKY TAACEMVAEMVESLQSVLALGHKRNSGYPA FLIFILLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME GESSPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRKKPLKALDTRFGKK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSRATIGALISHEKLLLQINPERELGSMS			1				QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
NILTSVTKÜRADKNAHNHIRLFEPLVIKÄLKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFULAQFEYTEVGGYRESSAIIPNIFF FLVILSYBEYHSKQIIGFBYIQLCDGIMASGR KAVTHAPALQPVHDLFVLRGTIKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQLADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRYEMILAANLQSSMAQLPMEEINRIGEY LQSSGLAQRHQRLYSTLDRFFRVLARMVDIL ACRRYEMLLAANLQSSMAQLPMEEINRIGEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMAFM MNSEFNLSLLAFLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYGSLPTLARLAQY LVVVSKLPSHLHPERKEDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLLYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKY TAACEMVAEMVESLQSVLALGHKRNSGYPA FLIFILLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME GESSPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRKKPLKALDTRFGKK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSRATIGALISHEKLLLQINPERELGSMS	1 :						SLRNMVOAEOENDTSGWFDVLOKVSTOLKT
YTTTTCVQLQKQVLDLLAQUVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYBRYHSKQIIGPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENBDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSLRPVDMLLRSMF VTPNTMASVSTVQLMSGILALIRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETISRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIIIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTIPALVLLWCQULLLVNHTIDTYRWW APVQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETTVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLAYHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPSLDLQAGGLOCCLALQPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLOGSVLALGHKRNSGYPA FLIFILLRNIISLAARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFETWATLLGVLVTQPLVMEG QEESSPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRKKPLKALDTRFGKK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
DSDQVFIGFVLKÖFEYEVGÖPRESEALIPNIFF FLVLLSYERYHSKGIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFVCQELGTLLMCLIIIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTONPKYI TAACEMVAEMVESLQSVLAIGHKRNSGVPA FLIFLLRNIISLARLPLVNSYTRVPPILVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKKEFIVR INTLGWTSRTQFFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLUXSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGKK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1						,
FLVLLSYERYHSKQIIGPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLQYHQVLEMFILVLQQ CHKENEDK WKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQKKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHAALVLLWCQILLVNHITDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIDJISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFFA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLIBEQIPLSLDLQGAGLDCCCLAQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTFKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPFRELGSMS	1 1						
KAVTHAIPALOPIVIDLFVLRGTINKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDEW KWRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPVLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHITTYCQELGTLLMCLIIIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTFKKRSLSSTKLLSPQMSGEEDSDLA AKLGMCNRETVRRGALIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTFTMLKKTLQCLEGI HLSQSGAVLTYQDRLLTFFRVLARMVDIL ACREVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLBEQPIESLDLQAGLDCCCLALQPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTFKAISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEEPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS				1			
LETOKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQMHI DSHEALGVLNTLFEILAPSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQUKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVGAQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHAGIPLSDDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTIPKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFFEETWATLLGVLVTQLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	ŀ						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEGDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDCHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISSGQKSALFEA ARREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALVQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPHTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGDFGTAFPEIPVEFLQEKEVFKEFTYR INTILGWTSRTGPETFWATLIGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEGDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDCHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISSGQKSALFEA ARREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALVQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPHTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGDFGTAFPEIPVEFLQEKEVFKEFTYR INTILGWTSRTGPETFWATLIGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							LETOKEVVVSMLLRLIOYHOVLEMFILVLOO
DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTTPYCQELGTLLMCLIIIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRVEMILAANIQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTOPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIYEGEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1						
VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTTYCQELGTLLMCLIIIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSFS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQULSPERRTNTFKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVSQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNIPEETTSRFLIQLVGILLEDIVT KQLKVEMSEQHTFYCOELGTLLMCLIIIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLVNHITDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA ARLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLVSLLDBFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTTNTPKAISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR NTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVILARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGIDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							,
KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVILWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HILWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGHQAIOSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSILAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTTKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS						ĺ	IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVILWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HILWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGHQAIOSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSILAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTTKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							EHSEGKOIKNLPEETFSRFLLQLVGILLEDIVT
GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPK YI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHI YQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	j .						•
ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKITLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS				1		1	
AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTOPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							· · · · · · · · · · · · · · · · · · ·
HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						AKLGMCNREIVRRGALILFCDYVCQNLHDSE
AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS]]					ļ ,	
HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	Į [
ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1					Ī	
PPVSSHPLÖGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSILAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS				ļ		i	LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	[ĺ	İ	
AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS]]						
AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1				ĺ		
LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPK YI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1		i				
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	į J				J	ļ	
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			l			1	LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS						Į	
EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVYQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			į į			l	
TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS]				ļ		1
WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		İ		1	į		
INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		i		1	1	1	FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			ļ	ł			
QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS			1	.			
VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	j 1				j		
PVPSLSPATTGALISHEKLLLQINPERELGSMS					l	.	
PVPSLSPATTGALISHEKLLLQINPERELGSMS	1 1		1		ŀ	į	LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
		l					
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			1	ļ			
							TIMES TOMO, WEST-SILI DISTRICT ADDITIONS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of, peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
						LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA SGETDSE
430	1780	Ā	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS DFLLIILKEILQKRSDLHLILMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKVLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
431	1781	A	3474	I	-	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEQ ID NO: of nucl- eotide seq- ucnce	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Argininc, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
		\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \				QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLIPPLLYQQLLHS SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS APRSRCVARPAARTGLPTPAPASSPAPAASPA PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP GAPPPRPAASPSPAASPALTASPPLP AASPSPAASPAPPAASPVLTASPPLP AASPSPAASPAVLTASPPLPAASPSPA ASPAPPAASPVLTASPPLPAASPSPA ASPAPPAASPVLTASPPLPAASPALAASPVHT ASPPVHVASPPVHTASPPVHVASPPVHVASPPVHV ASPPVHTASPPVHVASPPVHVASPPVHVASPP VHTASPPVHVASPPVHVASPPVHVAYPPVHV ASPPVHVASPPVHVASPPVHVAYPPVHV ASPPVHVASPPVHVASPPVHVAYPPVHV ASPPVHVASPPVHVASPPVHVAALP
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV RTKFGINMVTSRERGTTRLPKEG
435	1785		3529		3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS FPAFAVEDSQLVELTKQEIITIKLQGRYGCCRF LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC EWPLPWTYFILDGVFSGNAEQVQEYKEALEA VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT VDRVPMGKLPHMWGQSLYILGSLMAEGFLA PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK LYDIRKTIFTTFQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE DGTSLNSSILAALRKMQDGYFGGARVQTGKL SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT TCDLMSLVTKAKELHVQNVHMYLPTKLFQA SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ SGEVDFKALVLQLKETSSLQEQADILYMLYT MKGPDWNTELYNERSATVRELLTELYGKVG EIRHWGLIRYISGILRKKVEALDEACTDLLSH QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA SEGDMSISILTQEIMVYLAMYMRTQPGLFAE MFRLRIGLIIQVMATELAHSLRCSAEEATEGL MNLSPSAMKNLLHHILSGKEFGVERSVRPTD SNVSPAISHEIGAVGATKTERTGIMQLKSEIK QSPGTSMTPSSGSFPSAYDQQSSKDSRQQW QRRRLDGALNRVPVGFYQKVWKVLQKCH GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL NRVPQPEYRQLLVEAILVLTMLADIENHSIGS IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ
436	1786	A	3546	73	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion KGKGKTIRGI*TFKGRKGGTYQREHDANPLA PXSARSCWMRKG
437	1787	A	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG SASGCRYAAHPHWGLGGAAAAGGSWEPQPP RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA AQPGARARTSPPPASARNMAARPAATLAWSL LLLSSALLREGCRARFVAERDSEDDGEEPVVF PESPLQSPTVLVAVLARNAAHTLPHFLGCLER LDYPKSRMAIWAATDHNVDNTTEIFREWLK NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDQAQLDWELIYIGRKRMQVKEPEKA VPNVANLVEADYSYWTLGYVISLEGAQKLV GANPFGKMLPVDEFLPVMYNKHPVAEYKEY YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT TLG
439	1789	Α	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLSTLTGRSY GQPSLQDELKDNTTVFTRILDRLLDGYDNRL RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI DVFFRQSWKDERLKFKGPMTVLRLNNLMAS KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFPMD\AH ACPLKFGSYAYTRAEVVYEWTREPARSVVV AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	Α	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK FWEVISDEHGIDPTGTYHGDSDLQLERINVYY NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569		1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ FIEELITKWQKNDQELISDPLQQCFKKDEILDG QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG QDRLAVLLPGRHPCDCLGQKHKLINNCLICG RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS NKSQKLLKKLMSGVENSGKVDISTKDLLPH QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ľ		peptide	sequence	/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
				-		DDESDYFASDSNQWLSKLERETLQKREEELR
1		!				ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
}						SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
						VNPNMYQSPPQWVDHTGAASQKKAFRSSGF GLEFNSFOHOLRIODQEFQEGFDGGWCLSVH
İ						OPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
						TAKKPSPQEVSELQATYRLLRGKDVEFPNDY
						PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS
1		ļ				PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
						QGAKKGLMKQNKAV
442	1792	Α	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK
					ļ.	KTAGVKPYECTICGKAFMRLSSLTRHMRSHT
						AIRAI\EKPYKCKEC\GRAFSLSQILSK\HERSH
		1	ŀ			TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK PYECKOCGKALSCSSSLRVHERIHTGEKPYEC
		}				KOCGKAFSCSSSIRVHERTHTGEKPYACK\EC
		}				GKAFIS/TTSVLTHMITHNGDRPYKCKECGKA
		1				FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
		1				TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV
		ĺ			,	HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
						HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG
1						EKPYECKECAKTFISLENFRRHMITHTGDGPY
Ì		l.				KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK
						AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
· .						LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
						LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
]						QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
						PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL
						VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
						ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK
'''		1	55.0	207		LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
		ì				MTKVTLENFYSNLIAQHEEREMRQKKLEKV
	ĺ	ļ.				MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
		i				RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH
]				VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM
						MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL
			1			GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
		1				KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE
		1				TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
	['		l			KLCDWWSLGVIMYEMLIGYPPFCSETPQETY
		l				KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
		1				AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE
		[1			TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
	{					KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD
						GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK
}	1					NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
1						FCKLGVLVRHRISLFGNDATSIVNCLHILGQT
						LDARTVMKTGLESVKSALRAFLDNAAEDLE
]					KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI
	1					LTSLYALGTSKSIYVERQRSALGECLAAFAGA
	1]			FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP
	1		-			TNVEDVCPNIPSLEKLMEEIVELAESGIRYTQ
				·	·	· · · · · · · · · · · · · · · · · · ·

l l		1				SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML PRLVSNSWTQAILLPRPPKMLGLQV
		ļ				DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL
446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
1 1						FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN
		}				FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
[]		}				VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
} }	}	1			Į	RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI
		}		l i		NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF YNKSEDGDTPDMKCDDMLTCYMFHMYVGV
				Ì	}	GHY\NNFFFAAHLLDIAMGFKTLRTILSSVTH
				ļ		VKYQMWKLGVVFTDNSFLYLAWYMTMSVL
 			}	İ	}	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
	}			-	ŀ	LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG
]						PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
		.				KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL
}					1	FWKKIIAYQQKLLNYFARNFYNMRMLALFV AFAINFILLFYKVSTSSVVEGKELPTRSSSENA
				İ	. [KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
	ļ			ļ	Ì	KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
}						EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE
		·			<u> </u>	KKIKVAELLANMPDPTQDEVRGDGEEGERKP LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
						FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA
			·			LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI
						QGPRMAFFSILTVRSALFALRYNILTLMRMLS
					Į	VKESKRQFIFDVVNEGGEKEKMELFVNFCED TIFEMQLAAQISESDLNERSANKEESEKERPEE
						FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ
						VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP
						SCAETDENETLDYEEFVKRFHEPAKDIGFNVA
						ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD PDGKGVIFKRDFHKAMESHKHYTQSETEFLL
]			}			KDMVVMLLSMLEGNVVNGTIGKQMVDMLV
Į į						VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
						KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
						ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA
			l			VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVO
						FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
-						LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF
			Ì			CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ ARLHDRGAAEMVLQTISASKGETGPMVAAT
ł						LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
						YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ
						RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH
			{			DPEKTVERVLDIANVLFHLEQKSKRVGRRITY CLVEHPQRSKKAVWHKLLSKQRKRAVVACF
						NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
						NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS
1						ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
]			KEPNPEAEELFRMVAEVFIYWSKSHNFKREE ONFVVONEINNMSFLITDTKSKMSKAAVSDO
1.	•					LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL
		}				EKLKKKAATVVSEEDHLKAEARGDMSEAEL
1					[GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM
1			{			AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE
		<u> </u>	-	Sequence		MPHVMEVILPMLCSYMSRWWEHGPENNPER
1	Ì			peptide sequence	:	/=possible nucleotide deletion, \=possible nucleotide insertion
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
uence	20,,00		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	HISolcucine, K-Lysine, L-Leucine, M-Methionine, N-Asparagine, P-Proline,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,

SEQ II DNO: of mucle colde USN DNO: of mucle colde USN Document DNO: of mucle colde USN Document DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of mucle colde USN DNO: of public USN DNO: of public USN DNO: of public USN DNO: of public USN USN DNO: of public USN USN DNO: of public USN DNO: of publ	LOCO IE	1 650 35	114.	LCEC	D-1-42 4 3	David Control	[A
nuclectide sequence contested to USSN polylose puncte unice contested to the contested polylose puncte unice unice contested polylose puncte contested polylose puncte contested polylose puncte contested polylose puncte contested polylose puncte contested polylose punctes contested polylose	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			noa				
uence 9/496 09/496 ontsponding on the size of periodic residue of peptide requence peptide sequence			ł				
1914 ng to first amino acid residue of peptide sequence peptide			Į.				
amino acid residue of peptide residue of peptide residue of peptide sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period of the peri		uence	ļ				
residue of peptide pep	uence	ŀ	1	914			
Peptide	1	1		1	t e		
		1	ľ	i	t .	sequence	
1797 A 3598 1202 1070	ł	ł	ł	!			
GR*V*GAAASQGA/GSLRPRTTOFOHPGAWL QVWGAAAVCAGPM/MVAKARGPRAG-EP NSPWRSGVLAARAVGAGPWP*P*PGCS*ARG PSSSSAPGLASGPAAPLLQVHSSAGPLLCVT NGTLALGLK**AWGWGEWPKG FREGGGGREDFGGALKYNSRHEKVNGLEE GREELPVNNNKLYEKHGPGRWVVLAAVLIG LLJALLGIGFLVWHLQYDDVRVQKVFNGYM RITNENFVDAYENSNTEFVSLASKVKDALKL LLSGGVPUFLGPYHKESSAVTAFSEGSVATYWEE FSPQHLVEEAERVMAERVVMLPPRARSIKS FVYTSVVAPTDSKTVQRTQDNSCSPGLHAR GVELMRFTTGFTDSTVQRTQDNSCSPGLHAR GVELMRFTTGFTDSTVPSNTLTHISSQNVL LITLITITITERRIPGGEATFFQLPMSSCGGRL RAGGTTSNSPYYFGHARQVALROD ADSVLSLTRSFBLASCDERGRILVTV*NTTL SIMEPHALVQLCGTVPSNTLTHISSQNVL LITLITITITERRIPGGEATFFQLPMSSCGGRL CDGWADATCHISDELNCSCDAGHGFTCKNKF RAGGTTSNSPYYFGHAPROKORHOUGGSSIDSTOR EYYCGERSGPVTSNSNTVRFBSQSYTDT GFLASYLSVDSSDPCCGGFTCRTGRCIRKELR CCPLFWVCDSINDCGDNSDGGCSCPVQTTF RCSNGKCLSKSQQCGKBDCGGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDSBEKDCDCLGLSSTRQARVYGGTD ADEGEWPWQVSLHALGQGHICGASLISNNWL VSAHACVIDDRGFRYSDPTQWTACLHDQS QRSAPGVQERRLKRISHFFNDFTFDYDIALL LEKPRASTSMVTRICDLASSTRTQARVYGTD ADEGEWPWQVSLHALGQGHICGASLISNNWL TOWGHTQYGGTGALLLQKGERVNGTTCEN LLPQQITPRMMCVCGFLGASSTVANDFVG ANFOLLPHCLGFEDVYIVKNKPVLLVCKAV TOWGHTQYGGTGALLLQKGERVNGTTCEN LLPQQTTPRMMCVGFLGSGVDSCQGSGGPL SSVEADGRIFGAGGVVSWGDGCAQRNKPGVY TILLFRDWIKENTG AFTERDWIKENTG APPOLEPHAGEFEDAVSCSTTTUTUTGDA TSPEGETDLCNLANRVHSPPLKKLSKRHLKVST ASLTSVDTAGHBULVNGVAVGGRAAPPVG ANFOLLPHTLGFFEDVYIVKNKPVLLVCKAV SGSTLSSQKAYBLAYLRKNFEGDELAKEVSL EQGIVLPCRPPEGIPAE 449 1799 A 3618 2 613 FYSGSPWRMDGTERLEARRAGRLFWSSQ EMPRESTSSANSILTVTTPFCDBADVS SGTTKSQKAYBLAYLRKNFEGDELAKEVSL EQGIVLPCRPPEGIPAE APSLEGGGMDT.TCPPGVSPACCAGASWSIFG ADAAEVPGTTGHSQCGAMPHPEDEEPPGE PAAASPAGQGGPFTAGVSCSFTTUTTUTGDA TSPEGETDLCNLANRVHSPPLKKLSKRHLLXVST ASLTSVDTAGHBULVNDQLPDISISEEDKKKN LALLEEAKLVSERFT.TRGRKSRSSPCDSFSA VSPNLSFSASTSSSSSLTVTTPFFCDBADVS SHFIGEFNVFKGLADRAGCHORKVSQCGAAP RPPVKSKELBGKENFEDLQVFETTFKGLA AFGLKDFQGVVFFVFSELGGKSPLLANGSBCEP GPRLQKVLAKLPLAEEKERAGKAGGKLAK AFGLKDFQQVVFVFNKSKLTKLRKFFRESPLFGASP LTEKEVENVFVQLSSAFRNDSYTLESRRQAE ERNLLTEENTEKELEBFKASTISSASLWHICE ERSGLDVMPINSDVTLLSRLVFIRESPTGAPP LTEKEVENVFVQLSSAFRNDSYTLESRRQAE ERNLTEENTE		<u> </u>				<u></u>	
OVWGAAAVCAGPAM*/AVRAKRGPRAG*PEP	447	1797	A	3598	1202	1070	
A NSPWRSGYLAARAVGAGFWP-PPGCS-AG		i				ĺ	
PSSRSAPGLASGPAPILLQGVHSSAGPILCYI MGTLALGILEPY AWGWGWEWPKG MGTLALGILEPY AWGWGWEWPKG MGTLALGILEPY AWGWGWEWPKG FRRKGGGGREFGAGILKYNSRIEEKYNGLEG GVEFLPYNNYKEKHEPGRWVLAAVLIG LLVLLGIGFLVWHLQYRDVRVQKVFNGYM RITMENFVDAYENSNSTEFYSLASKVKDALL LYSGYPFLGYPHKESAVTARSSGSVIAYYWSE FSIRQHLVEEAERWAEERVVMLPPRARSIS. FVTSVVAFFDSKTYQKTQDNSCSFGLHAR GVELMEPTTGPPDSPYPAHARCQWALRGD ABSVLS.TFRSTDLASCDERGRHLWITVYNTUL SPMEPHAU VQL.GGTYPPSYNLTHSSSOQUL LITLITITIERHRHFOFASTGFAGIKEL CDGWADCTDHSDELNCSCOAGHGFTCNKW CRYCGERSQFVVTSNSNKITVRHSDQSYTD GFLASTSYSSDPCPGGFTCRTGREIKEL CDGWADCTDHSDELNCSCOAGHGFTCNK CRYCGERSQFVVTSNSNKITVRHSDQSYTD GFLASTSYSSDPCPGGFTCRTGREIKEL CDGWADCTDHSDELNCSCOAGHGFTCNK CRYCGERSQFVVTSNSNKITVRHSDQSYTD GFLASTSYSSDPCPGGFTCRTGREIKEL CDGWADCTDHSDELNCSCOAGHGFTCNK CRYCGERSQFVVTSNSNKITVRHSDQSYTD GFLASTSYSSDPCPGGFTCRTGREIKEL CDGWADCTDHSDELNCSCOAGHGFTCNK CRYCGERSQFVVTSNSNKITVRHSDQSYTD GFLASTSYSSPVTGLDGCSDGSDEASCP KVNVVTCTKHTYRCLNGLLSGFTRQARVVGGTL ADEGEWPWOYSLHALGQGHICKTGREIKEL CDGWADCTDHSDELNCSCOAGHGFTCNK CRYCGERSQFVTTSNSNLTVRHSDQSTD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ADEGEWPWOYSLHALGQGHCYARTALLOLD ANTON	1		1	ļ	·	ļ	§ 7
1798 A 3604 3115 557 FRRKGGGGPRFGAULYNSRIEVNGLEE GVEFLPVINVLKVEKHOPRWVLAAVLIG ELLVYLLGIGFLWHLQYRDVRQKVFNGYM RITNENFVDAYENSNSTEFVSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKVKDALKL LYSGYPFLGFYHKESAVTAFSEGSVIAYVWENSKEFFYSLASKDEGGREGFRH.VITVYTHL SPMEPHALVQLCGTYPPSYNLTHSISONVL LITLITINTERHPGFEATFFQLPRMSSCGGR RKAQGFTNSPYTGHYPRIDACTWNIEVPRN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVYENN QHVKVPKFFYLLEFQVPAGTCFKDVFCNN QHVKVPKFFYLLEFQVFAGTCFKDVFCNN QHVKVPKFFYLLEFQVFAGTCFKDVFCNN QHVKVPKFFYLLEFQVFAGTCFKDVFCNN QHVKVPKFFYLLEFQLASKGPECDOK QHSAGFQFTCATGCTRELEFT QHVKVPTCTKTTYRCLLMGLASKGPECDOK QHSAGFQVFCTCHASKGPTCHASKGPT QHVKVPTCTKTTYRCLMGLLSKGMPECDOK QHSAGFQVFCTCHASKGPTCHASKGPT QHVKVPTCTKTTYRCLMGLLSKGMPECDOK QHSAGFQVFCTCHASKGPTCHASKGPT QHVKVPTCTKTTYRCLMGLSKGMPECDOK QHSAGFQVFCTCHASKGPT	l	!	İ		{	i	NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG
448 1798 A 3604 3115 557 FRRKGGGGPKDFGAGLKYNSRIEKVNGLEG GVEPLPVNNYKKVEKHOPGRWVLAAVLIG LLLVLLGIGFLWHLQYBQVEVPGVPNYKK RTNENPVDAVENSNSTEFVSLASKVKDALKL LYSGVPFLGPYHKESAVTAFSGGSVIAYYWSE FSIRQHLVEEAERWAEERVVMLPPRARSLS FVYSVVAFFIDSKTVQRTQDNSCSFGLHAR GVELMEFTTRGFPDSPYPAHARCQWALRGI ADSVLSLTFRSTDLASCDERGRHLWTVYNTIL SPMEPHAU VQL.GGTYPPSYNLTHSUSGOVL LITLITINTERHEPOLASCDERGRHLWTVYNTIL SPMEPHAU VQL.GGTYPPSYNLTHSUSGOVL LITLITINTERHEPOLASCDERGRHLWTVYNTIL SPMEPHAU VQL.GGTYPPSYNLTHSUSGOVL LITLITINTERHEPOLASCDERGRHVSVYDENG EXYCGERSOPVYTSNSNLTVRHSDQSYDTOT GFLASYLSYNSSDPCPGGFTCRGRVLERELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSBQGCSCPAQTIT RCNGKCLSKSQQCNKDDCDGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGPPEDDK EEXYCGERSOPVYTSNSNLTVRHSDQSYDTOT RCNGKCLSKSQQCNKDDCDGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGPPEDDK EEXYCGERSOPVYTSNSNLTVRHSDQSYDTOT RCNGKCLSKSQQCNKDDCDGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGPPEDDK EEXYCGERSOPVSDTGOMTAFLGLHDQS QRSAFGVQERLKRIISHFFFNDFTDVDIALL LELEKRAYSSNVPGCLOGASTPFTDVDTALL LELEKRAYSSNVPGCLOGASTPTTOTAPLGLHDQS QRSAFGVQERLKRIISHFFNDFTDVDTALL LELEKRAYSSNVPGCLSGCDYSCOGBGGPL SVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKSNTGV LIPLQTTPRMCVGCJSGCPSCOGBGGPL SVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKSNTGV AND TRLPLFRDW	ł	ł	1	1 .	}		
GVEFLPVNNYKKVEKHGPGRWVVLAVLIG LLLVLLGIGI-LWHLQVRDVRVQKVPNGYM RITNENFVDAVENSNSTEVSLASKVKDALKL LYSGVPFLGPVHKESAVTASKSVKDALKL LYSGVPFLGPVHKESAVTASGSVLAYYWSE FSIPQHLVEEAERVMASERVVMLPPRASILKS PVVTSVVAPPTDSKTVQRTQDNSCSFGLHAR GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGHLWTVYNTU SPMEPHALVQLGSTYPPSYNLTHSSSQNVL LITILITIMTERRHPGWEATFFQLPRMSSCGGRL RKAQGTINSPYYPGHYPPNISHDOSYTDT GFLASTLSVDSSDPCTQFTCRTGRCIRELR CDGWADCTDHSDELNCSLOGGHLCRKELR CDGWADCTDHSDELNCSLOGGHCASGRLS CDGWADCTDHSDELNCSLOGGHCASCLSPNWL COGWADCTDHSDELNCSLOGGHCASCLSPNWL COGWADCTDHSDELNCSLOGGHCASCLSPNWL VSAAHCYIDDRGFRYSDPTQWTAFLLGLIQG QRSAFQVQERKLKRISHPFNFFTPDTDIALL ELEKFAEVSSMYRPICLPDASHVPPAGKAIW TGWCHTQYGGTGALLUKGGARVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLLGLIQG QRSAFQVQERKLKRISHPFNFFTPDTFDDIALL ELEKFAEVSSMYRPICLPDASHVPPAGKAIW TGWCHTQYGGTGALLUKGGARVGGTG LLPQGTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRITQAGVVSWGDGAGVNSGTGANGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSIMAGRQHGWSAQQSATVANPVPG ANPDLIPHI-GRGPDVYIVKNPLLVCKAV PATQIFFRCNGCWVRQVDHVIERSTDGSSGLP TMSVRJNSSGSGPTAGVSCCGVANSHVAVLLVCKAV PATQIFFRCNGCWVRQVDHVIERSTDGSSGLP TMSVRJNSPASSFTVAVCQCVAWS SGTTKSQK AVTRIAVLRNFEQEPLAKEVSL EQGVLPCRPPEGIPPAS 450 1800 A 3620 1 2676 MEPSLOGGMDLTCPFGVSPACGAQASWSIFG ADAAEVYGTRGHSQGAAMPHIFEDEEPFGE PQAAQSPAGQQFTAGVSCSCPTPTIVLTGDA TSPEGETINKLANR VHSPHKRLSHRHLKVST ASLTSVPPAGHIDLVNDQLPSTEPTGELAKEVSL EQGVLPCRPPEGIPPAS VSNISPSASSFTSSRSSILTYPTPFEGDEADVS SPIPGEPNVPKGLADRKONDORKVSGGRLAP PPPVEKSKELAHCKSTSPDPLAGLSWDSGPE PGPPLQKVLAKLPLAEEKRFAGKAGGKLAK PVTNSSGKMALNSPQPGPVSELGGQLLKTG WEGSPLPRSPTDQAGVGVGPTTPTFGGLAA PVTNSSGKMALNSPQPGPVSELGGQLLKTG WEGSPLPRSPTDQAGVGVGPTTPTFGGLAA PVTNSSGKMALNSPQPGPVSESELGKQLLKTG WEGSPLPRSPTDAAGCGPAGGRPAGEP GPPCAKSCLELEKKRSTSRSSNSUHCHE EKSKGLDVMRPHISDVLLRILRKRHSSLFGSAPP LTEKEVENVFVQLSSAFRNDSYTLESKINQAE RERNITEENTEKELENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLR		!]	ł .	ļ		
GVEFLPVNNYKKVEKHGPGRWVVLAVLIG LLLVLLGIGI-LWHLQVRDVRVQKVPNGYM RITNENFVDAVENSNSTEVSLASKVKDALKL LYSGVPFLGPVHKESAVTASKSVKDALKL LYSGVPFLGPVHKESAVTASGSVLAYYWSE FSIPQHLVEEAERVMASERVVMLPPRASILKS PVVTSVVAPPTDSKTVQRTQDNSCSFGLHAR GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGHLWTVYNTU SPMEPHALVQLGSTYPPSYNLTHSSSQNVL LITILITIMTERRHPGWEATFFQLPRMSSCGGRL RKAQGTINSPYYPGHYPPNISHDOSYTDT GFLASTLSVDSSDPCTQFTCRTGRCIRELR CDGWADCTDHSDELNCSLOGGHLCRKELR CDGWADCTDHSDELNCSLOGGHCASGRLS CDGWADCTDHSDELNCSLOGGHCASCLSPNWL COGWADCTDHSDELNCSLOGGHCASCLSPNWL COGWADCTDHSDELNCSLOGGHCASCLSPNWL VSAAHCYIDDRGFRYSDPTQWTAFLLGLIQG QRSAFQVQERKLKRISHPFNFFTPDTDIALL ELEKFAEVSSMYRPICLPDASHVPPAGKAIW TGWCHTQYGGTGALLUKGGARVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLLGLIQG QRSAFQVQERKLKRISHPFNFFTPDTFDDIALL ELEKFAEVSSMYRPICLPDASHVPPAGKAIW TGWCHTQYGGTGALLUKGGARVGGTG LLPQGTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRITQAGVVSWGDGAGVNSGTGANGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSIMAGRQHGWSAQQSATVANPVPG ANPDLIPHI-GRGPDVYIVKNPLLVCKAV PATQIFFRCNGCWVRQVDHVIERSTDGSSGLP TMSVRJNSSGSGPTAGVSCCGVANSHVAVLLVCKAV PATQIFFRCNGCWVRQVDHVIERSTDGSSGLP TMSVRJNSPASSFTVAVCQCVAWS SGTTKSQK AVTRIAVLRNFEQEPLAKEVSL EQGVLPCRPPEGIPPAS 450 1800 A 3620 1 2676 MEPSLOGGMDLTCPFGVSPACGAQASWSIFG ADAAEVYGTRGHSQGAAMPHIFEDEEPFGE PQAAQSPAGQQFTAGVSCSCPTPTIVLTGDA TSPEGETINKLANR VHSPHKRLSHRHLKVST ASLTSVPPAGHIDLVNDQLPSTEPTGELAKEVSL EQGVLPCRPPEGIPPAS VSNISPSASSFTSSRSSILTYPTPFEGDEADVS SPIPGEPNVPKGLADRKONDORKVSGGRLAP PPPVEKSKELAHCKSTSPDPLAGLSWDSGPE PGPPLQKVLAKLPLAEEKRFAGKAGGKLAK PVTNSSGKMALNSPQPGPVSELGGQLLKTG WEGSPLPRSPTDQAGVGVGPTTPTFGGLAA PVTNSSGKMALNSPQPGPVSELGGQLLKTG WEGSPLPRSPTDQAGVGVGPTTPTFGGLAA PVTNSSGKMALNSPQPGPVSESELGKQLLKTG WEGSPLPRSPTDAAGCGPAGGRPAGEP GPPCAKSCLELEKKRSTSRSSNSUHCHE EKSKGLDVMRPHISDVLLRILRKRHSSLFGSAPP LTEKEVENVFVQLSSAFRNDSYTLESKINQAE RERNITEENTEKELENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLRILRKLEENFKASKSLWHHCE EKSKGLDVMRPHISDVLLR	448	1798	Α	3604	3115	557	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE
LLLVLIGIGFI.VWH.QVDRVRVQKVPNGYM RITNENFUDA VENSINSTEPS LASKVKDALKL LYSGVPFLGPYHKESAVTAFSEGSVIAYYWEE FSIPQHLVEEAERVMALERVALPPRARSLKS FVVTSVVAPPTDSKTVQRTQDNSCSFGLHAR GVELMRTTIFGFDSFYPAHARCQWALRGD ADSVLSLIFRSFDLASCDERGRHL.VITVYNTL LITLITNTERRHPGGFATTFGLPRMSSCGGRL RKAQGTFNSPYYPGHYPNIDCTWNIEVPNN QHYWRFKFFYLLEFQFPAGTCFKDYVENG EKYCGERSQPVVTSNSNLITVESISQNYL LITLITNTERRHPGGFATTFGLOFRMSSCGGRL RKAQGTFNSPYYPGHYPNIDCTWNIEVPNN QHYWRFKFFYLLEFQFPAGTCFKDYVENG EKYCGERSQPVVTSNSNLITVEFISDQSYTDT GFLAEVLSYNSDPCTQGFTCRGTRELER CDGWADCTDHSDEL NCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPPAQTF RCSNGKCLSKSQQCMGKDDCDGOSDEASCP KVNVVTCTRHTYRCLNGLCLSKGMPECDOK EDCSDSDSENDCGGLSSFNQRVVGGTD ADEGEWPWQVSLHALGQGHIGGASLISPNUL VSAAHCTIDDAGFRYSDPTQHAFLGLIHDQS QRSAPGVQERLKRIISHFFNDFTFDVDIALL ELEKPAEYSSMYPICLPDASCQGDSGGPY LSPCAMWAWCVGFLSGGVAWFOCYDT TRUPLFRDWIKENTGV TRUPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRWDGSTERLEARRPAGRLPWSSRQ EMTRRSLMAGRQHGWSAQQSATVANIPYPG ANPDLIPHICLGFEDYVIVKNEPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSCLP TMSVRNIVSRQVEKVFGLEFBYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFGQEPLAKEVSL EQGIVLPCRPPGIPPAE 450 1800 A 3620 1 2676 MEPSLQGGMDLTCPFGVSPACGAQASWSIFG AASLTSVDPAGHIDLVNDQL-DISISEEDKKN LALLEEAKLVSERFLTRRGRKSRSSFORDSFA VSSNLSPASSPTSSSNSLTVTPTPEGDEADVS SPHOGEPNVPKGLADRKQNDQRKVSGGRAAP RPPPVEKSELBIFCLSFNLSSDSPA VSSNLSPASSPTSSSSNSLTVTPTPEGDEADVS SPHOGEPNVPKGLADRKQNDQRKVSGGRAAP RPPPVEKSELBICKENEDPLYPTTTPTICA PVTNSSGMALNSPQGFVSELGKQLLKTG WEGSPLPRSPTQDAAGVGPPAGGRAAGRALAK APGLKDFQUQVPVRMQKLTKLREEHILMRN QNLVGLKLPDLSRAAQEKCLPSELSFALEGE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVMRPISDVLLRILKRHSLLEFKARFAREE EKKSGLDVARPHISDVLLRILKRHSLLEFKARFAREE EKKSGLDVARPHISDVLLRILKRHSLLEFKARFAREE EKKSGLDVARPHISDVLLRILKRHSLLEFKARFAREE EKKSGLDVARPHISDVLLRILKRHSLLFGSAPP LTEKEVENVFYQLSSAFRNDSYTLESKINQAE RERNUTEENTERELEFKARSASKWHHCE	1	!	1			•	
RITNENFVDAVENSNSTEFVSLASKVKDALKL LYSGVPFLGPVHESAVATASSILAS FVVTSVAPPTDSKTVQRTQDINSCSFGLHAR GVELMRTTIFGFDSYPAHARCQWALRGD ADSVLSLIFRSFDLASCDERGRHLWTVYNTU SPMEPHALVQLGTYPPSYNLTHSISQNVL LITLITNTERRHPGVEATFFQLPRMSSCGGRL RKAQGTINSPYPOHYPNIDHSDORYDTI GFLAEYLSYDSSDPCPGQFTCRTGRCTRKELR CDGWADCTDHSDELNCSCLOAGHQFTCKNF CKPLFWCDSJLDCSDNSDEGGCSCPAQTIT RCSNGKCLSKSQCQNGKDDCGDGSDEASCP KVNVTCTKHTYRCLNGLCJKKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGGGHASLISPNWL VSAAHCYDDRGFRYSDPTQWTAFLGLHOQS QRSAPOVQERRLKRISHPFTFDYDIALL ELEKAAEYSSMYRPICLPASHYPAGKAHW TGWGHTQYGGTGALLLQKGERVNQTTCEN LLPQGTPRMMCVGFLSGGYDSCQGBSGGPL SSVEADGRIPQAGVVSWGDGCAQRNKRYY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSIMAGRQHGWSAQQSATVANPVG ANPOLLPHFLGEPEDVYIVKNRPVLLVCKAV PATQIFFKCNGEWVRQVDHVIKSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTITKSQKAYIRLAYLRKNPEQEPLAKEVSL EQGIVLPCRPEGIPAE 450 1800 A 3620 1 2676 MEPSLQGGMDLTCPFGVSPACGAQSWSWFG ADAAPVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSTITIVLTGDA TSPEGGTIKKLANRNHSPHKRISHRHLKVSI ASLTSVDPAGHIDLVNDQLPDISISEDKKN LALLEEAKLVSERFLTRRGRSSSPCDSFSA VSPNLSPSASPTSSRSNSLTVTPPEGDEADVS SPHPGEFNVFKGLADKKQNDQRVSQGGRAGF PGPAQSPAGQQGPTAGSVCSTITIVLTGDA TSPEGGTIKKLANRNHSPHKRISHRHLKVSI ASLTSVDPAGHIDLVNDQLPDISISEDKKN LALLEEAKLVSERFLTRRGRSSSSPCDSFSA VSPNLSPSASPTSSRSNSLTVTPPEGDEADVS SPHPGEFNVFKGLADKKQNDQRVSQGGRAGF PGPPQKSKCLALROKAGNGLKAK PVTNSSGKMALNSPQPGPVESELGKLLKTG WEGSPLPRSTPDAAGVGUPPTTPRGLA PVTNSSGKMALNSPQPGPVESELGKLLKTG WEGSPLPRSTPDAAGVGUPPTTPRGLA PVTNSSGKMALNSPQPGPVESELGKLLKTG WEGSPLPRSTPDAAGVGUPPTTPRGLA PVTNSSGKMALNSPQPGPVESELGKLLKTG WEGSPLPRSTPDAAGVGUPPTTPRGLA PVTNSSGKMALNSPQPGPVESELGKLKTG WEGSPLPRSTPDAAGVGUPPTTPRGLA PVTNSSGKMALNSPQPGPVESELGKLANGE EEKSGLDVMMPIISDVLLRKLRWHRSLPGSAP BGPPGCKVLAKLLPLAEEEERFAGAGGKLAK APGLKDFQQVQPVMQKLTKLREHHILMN QNLVGLKLPDLESAAEQEKGLPSELSPAIEEE EKSGLDVMPNISDVLLRKLRHWRSLPGSAP LTEKEVENVFQLSSAFRNDSYTLESSINQAE RERNITEENTEKELENFKASTSSASLWHHCE		ł	1			[
LYSGVPFLGPYHESAYATASSGGSULAYYWSE FSRPQHLVERARWAMERRVVM.PPRARSILKS FVVTSVVAPPTDSKTYQRTODNSCSFGLHAR GVELMRPTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFOLASCDERGHL.VITVYNTL SPMEPHALVQLCGTYPRSYNLTHSSSGNVL LITILITMERRHPGWEATFRQJPRMSSCGGRL RKAQGTFNSPYYPCHYPPNDDCTWNIEVPNN QHVXVRRFYPLLEFQYPAGTCPKDYVENG EKYCGRSGPVVTSNSKITVRFHSDQSYTDT GFLAEALSYDSSDPCTQGFTGRGIRELER CDGWADCTDHSDELNCSCDAGHGFTCRKFK CKPLFWVCDSLDNDCGSNDGGCSPAQTF RCSNGKCLSKSQQCQGKDDDCGDGSDASCP KVNVVTCTKHTYRCLNGLCLSKGNPECOK EDCSOSDEKDDCDGALSFRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNUL VSAAHCYIDDRGFRYSDPTQHAFLGLIHDQS QRSAPCVQERRLKRISHPFFNDFTFDVDIALL ELEKPAEYSSMYPICLPDASMYPGCAMPAGCAG LIPQOTPRMMCVGFLSGQVSQGGSGGPL SSVEADGRFQAGVVSWCDGCAGRNKRGVY TRIPLFRDWIKENTGV TRIPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSLMAGRQHGWSAQQSATVANPVG ANPDLIPHGLEFEDYVIKNRPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVCVEKVGLEFWCQCVAWS SSGTTKSQKAYIRIAYLRRNFEQEPLAKEVSL EQGIVLPCRPPGIPPAE 450 1800 A 3620 1 2676 MEPSLQGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTTGHSQQCAAMPILEDEEPDGG PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGGTLKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLFDSISEEDKKRN LALLEEAKLVSERFLTRRGRKSRSSFOGSPSA VSRNLSPSASPTSSRSSNLTVTPPEGDEADVS SPIPOGEPNVPKGLADRKQNDQRKVSQGRLAP PRPPVEKSKELBICKENEDPLQYFETTPRGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTDQAGVGVPRAGAGGRAAGE PGPRLQKVLAKLPLAEERERAGAGGKLAK APGLKDFQUQVPVRMQKLTKLREEHILMRN QNLVGLKLPDLSRAAQEKGLPSLSSRSSLWHHCE EKKSGLDVMRPISSULLRILKRYHRSLEGSAPP LITEKEVENVFVQLSSAFRNDSYTLESKINVAE REENVLTEENTEKELLEFKKASTSLWHHCE EKKSGLDVMRPISSULLRILKRYHRSLEGSAPP LITEKEVENVFVQLSSAFRNDSYTLESKINVAE REENVLTEENTEKELLEFKKASTSLWHHCE EKKSGLDVMRPISSULLRILKRYHRSLEGSAPP LITEKEVENVFVQLSSAFRNDSYTLESKINVAE REENVLTEENTEKELLEFKKASTSLWHHCE EKKSGLDVARPHISSULLRILKRIPSLEGSAPP LITEKEVENVFVQLSSAFRNDSYTLESKINVAE REENVLTEENTEKELEFKKASTSLWHHCE EKKSGLDVARPHSSULLRILKRIELBERASRASLWHHCE EKKSGLDVARPHSSULLRILKRILEREHILMRN QNLVGLKLPDLSRAARQEKGLFSLESSASLWHHCE EKKSGLDVARPHSSULVLRIKGLEFTENGSVASSLWHHCE EKKSGLDVARPHSSULVLRIKGLEFTENGSVASSLWHHCE EKKSGLDVARPHSSULVLRIKGLEFTENGSVASS	1	į	1]	1
FSIPOHLVEEAREVMAEERVVMLPPRARSLKS FVVTSVVAPPTDSKTVQRTODSSTGULARG ADSVLSLTFRSFDLASCDERGHIL-VITYVTVTU SPMEPHAL VQLGTYPPSYNLTHSISSQNVL LITILITINTERRHPGVEATFFQLPRMSSCGGRIL RKAQGTINSPYPCHYPPNISPDGSYTDT GFLAEVLSYDSSDPCPQFTCRTGRCIRELR CDGWADCTDHDBLI-NCSLOGHQFTCKNSK- CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQCORGKDDCGDGSDEASCP KVNVTCTKHTYRCLNGLCJKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALOGGHCSALISPNUL VSAAHCYIDDRGFRYSDPTQWTAFLGLHOQS QRSAPGVQERRLKRIISHPFTDFTDVIALL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYGGTGALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPPAGKAHW TGWGHTQYSGGTALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPAGKAHW TGWGHTQYSGGTALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPAGKAHW TGWGHTQYSGGTALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPAGKAHW TGWGHTQYSGGTALLIG VGGAGSLISPNUL ELEKPAEYSSMYRPICLPDASHVPAGKAHW PATQIFFKCNGGWVRQVDHVIRRSTDGSSGLP PQAAQSPAGQQGPTAGSVSCSTTITVLTGDA TSSFGGTIKNLANNHSPHKRISHHLIKVST ASLTSVDPAGHIDL-VNDQLPDISIEEDEPPGG PQAAQSPAGQQGPTAGSVSCSTTITVLTGDA TSSFGGTIKNLANNHSPHKRISHHLIKVST ASLTSVDPAGHIDL-VNDQLPDISIEEDEPGG PQAAQSPAGQQGPTAGSVSCSTTITVLTGDA TSSFGGTIKNLANNHSPHKRISHHLIKVST ASLTSVDPAGHIDL-VNDQLPDISIEEDEPGG PQAQSPAGQQGPTAGSVSCSTTITVLTGDA TSSFGGTIKNLANNHSPHKRISHHLIKVST ASLTSVDPAGHIDL-VNDQLPDISIEEDEPGG PQPAQSPAGQQGPTAGSVGCAFGAGGKLAK PVTNSSGMALNSPQPGPVESELGKQLLKTG WGGSPLPRSTTPQDAAGVGPGAGGGAGGAG PGFGGTAGVACKALLIG LEFERAEGAGGKLAK PUTNSGGMALNSPQGAGGGTAGGE PGFRCLKVLAKLILAEEHIL MRN QNLVGLKLPDLSEAAEQKGLDPESLSPAIEEE ESKSGLDWMPIGSDVLLRRLR	1 .	ļ	}	}		ļ	
FVYTSVVARFTDISKTVQRTQDNSCSFGLHAR GVELMRFTTGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLWTVYNTUL SPMEPHALVQLCGTYPFSYNLTFHSSQNVL LITLITITERRHPGIFFATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPFNIDCTWISEVPNN QHYKVRFKFFYLLEBQVAGTCPKDYVENG EKYCGERSQPVYTSNSNKITVRHSDQSYTDT GFLASYLSYDSSDPCFQCTGATGCRTRELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTIF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSBERDCDCGGRTRQARVVGGTD ADEGGWPWQVSLHALGQGHICGASLISPNWL VSAAHCYDDGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLLRIISHPFFNDFTFDVDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLQKGERVINQTTCEN LLPQQTTRMMCVGFLSGGVDSCQDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEFDVYIVNNKPLVLCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQLVAWS SSGTTKSQRAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGMDLTCPFPGVSPACGAQASWIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSTTIVLTGDA TSPEGETIDKALANRVHSPKLSIRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPFDEDEADVS SPIFIGEPNYPKGLADRKQHDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PYTNSSGKMALNSPQFPVESLIGKQLLKTG WGGSPLRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSKQLLKTG WGGSPLRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSKQLLKTG WGGSPLRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSKQLLKTG WGGSPLRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSKDLLTG WGGSPLRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSKDLSKTHLKVST ASLTSVDPAGHIDLVNDQLPDISISEDKKKN QNLVGLKLVSERFLTRRGKSRSSPGDSPSA VSPNLSFRASPTTOAAGVGPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSSLSPAGEAGGKLAK APGLKDPQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLVBLSLARAFGGKAGGKLAK APGLKDPQIQVGPVRMQKLTKLREEHILMRN QNLVGLKLVBLSLARAFGGKAGGKLAK APGLKDPQIQVGPVRMQKLTKLREEHILMRN QNLVGLKLVBLSLARAFGGKAGGKLAK APGLKDPQIQVGPVRMQKLTKLREEHILMRN QNLVGLKLVBLSLARAFGGKAGGGCLAR RERNLTEERTEKELENFKASSITSASSLWHHCE	i	ļ		!		į	1
GVELMRFITGSPDASYNAHARCOWALROD ADSVLSLTRSFDIASCRENL-WITVYNTU SPMEPHALL VQLCGTYPPSYNLTFHSISQNVL LITLITHTERRHPGFEATFFQLPRMSSGGRIL RKAQGTFINSPYYPGHYPBIDCTWNIEVPNN QHVKVRFKFFYLLEPGYPAGTCPKDYENG EKYGGERSQPVYSISNKITVRFHSIOGSYDT GFLAEYLSYDSSDPCPGQFTCGTGCRELER CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTI RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVGGTD ADEGGWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPIQWTAFLGLHDQS QRSAPGVQERRLKRIISHFFFNDFTDDIALL ELEKPAEYSSMVEPICLPDASHVFPAGKAIWV TGWGHTQVGGTGALLLQKGEIRWNQTTCEN LLPQQTTPRMCVGFLSGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPFFRDWIKENTGV 449 1799 A 3618 2 613 FVSSSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHWSQQQSATVANPVPG ANPDLLPHFLGFEDVYLVRKPVLLVCKAV PATQIFFKCNGEWVRQVDHVERSTDGSSGLP TMEVRINVSRQVEKVFQVDHVERSTDGSSGLP TMEVRINVSRQVEKVFQUEWVRQVDHVERSTDGSSGLP TMEVRINVSRQVEKVFQUEWVRQVDHVERSTDGSSGLP TMEVRINVSRQQVEKVFQNDHVENSTDGSGLP TMEVRINVSRQQVEKVFQNDHVENSTDGSPA TMEVRINVSRQVEKVFQNDHVENSTDGSPA TMEVRINVSRQVEKVFQNDHVTNTT TMETTAGAT TMETTAGATTAGAT TMETTAGATTA	l .	i		İ			
ADSVLSI.TRRSPOLASCDERGRHI.VITVYNTUL SPMEPHALV QULGTYPSWLTFHSISQNIVL LITLITNTERRHPGVFEATFFQLPRMSSCGGRL RKAGGTFNSPY YPGHYPFNIDCT WINEVPNN QHYKVRRKFYTLLEBOYDGYTDT GFLASYLSYDSSDCPGGPTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSEGGSCPWGFT RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSBERDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYTDDGRFRYSDPWAFLGLHDQS QRSAPGVQERLKRIISHPFNDFTFDVDIALL ELEKPASYSSMVZPRICLPMAFLGLHDQS QRSAPGVQERLKRIISHPFNDFTFDVDIALL ELEKPASYSSMVZPRICLPMAFLGAHDVA TGWGHTQYGGTOALILQKGEIRVNQTTCEN LLPQQITPRMMCVGFLGGVDSCQDBSGPL SSVEADGRIFQAGVVSWGDGCAQRNKRGVY TRLPLFRDWIKENTGV FYSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGFPEDVYIVKNFVLLVCKAV PATQIFFRCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLESWCQCVAWS SSGTITKSQKAYIRIAYLRKHFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASWSFG ADAAEVPGTRGHSQQEAAMPHIFEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTTIVLTGDA TSPEGETDIKNLANRVHSPIKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASFTSSRSNSLTVPTFPEDBEADVS SPFIGEPNYFKGLARKAGNGGKLAK APGLKUPSRFTQBAGVGVPASQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTFKCLA PVTNSSGKMALNSPQFGVVESELGKLLKTG WGGSPLRSSTTQBAGCVGPASQGRGRAGE MGPEAGSKAELPPTVSRPPLLRGLSWDSGFEE PGFRLQKVLAKLPLAEEERAGKAGGKLAK APGLKDPQIQVOPVRNQKLTKLREHILMRN QNI.VGLKLYGLSAFRAGKAGGKLAK APGLKDPQIQVOPVRNQKLTKLREHILMRN QNI.VGLKLYBLSAAEAGCKLAKTAGREHILMRN QNI.VGLKLYBLSAAEAGCKLAKTAGREHILMRN QNI.VGLKLYBLSAAEAGCKLAKTAG PGFRLQKVLAKLPLAEEERAGKAGGKLAK APGLKDPQIQVOPVRNQKLTKLREHILMRN QNI.VGLKLYBLSAAEAGCKLAKTAGSRGSAPP LTSKEVENVFVQLSSAFRNDSYTLESRIQAE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTSKEVENVFVQLSSAFRNDSYTLESRIQAE RERNITEERITEKELENFKASITSSASS.WHICE	1 '	Į	ı			i i	
SPMEPHAUL VOLCGTYPPSYNT, FIRISSO, ONL LITLITNTERRIPGEATFOL PRMSCGGR. RKAQGTFNSPYYPGIPT PROTECTION ON CHIVKYREFFYLLER GYPAGTCPKDY VEING EKYCGERSQFVYTSNSNKITVRHSDSOSYTDT GFLAEYLSYDSSDPCPGOFTCRTGRCIRKELR CDGWADACTHISDEL NECOAGHOPTCKINKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGGSDEASCP KYNVVYTCTKHTYRCLIGLELSKGNPECDGK EDCSDGSSDEKDCDCGLRSTTRQARVGGTD ADEGEWPWQVSLHALGGGHICGASLISPNWL VSAAHCYDDRGFRYSDPTQWTAFLGLHDQS QRSAFGVQERRLKRIISPFPNDFTFDYDIALL ELERPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLQKGEIRYNQTTCEN LLPQQITTPRIMMCVGFLSGGVDSCQGBSGPL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPLFRDWIKENTG SEGVEN VYTRLPHTGGEPDVIYVKNRPVLLVCKAV PATGIFFKCNGEWVRQVDHVIERSTDGSSGLP TMCVRINNSQQVEK VFGLEEYWCQCVAWS SEGTTKSQKAYIRAYLRKNFEQEPLAKEVSL EQGIVLPCRPPGVSPACGAQASWSIFG ADAAEVPGTRGHSQGEAAMPHIPEDEEPPGE PQAAGSPAGQQGPTPGVSPACGAQASWSIFG ADAAEVPGTRGHSQGEAAMPHIPEDEEPPGE PQAAGSPAGQQGPTGVSCSPPTTULTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKKSRSSFGDSPSA VSPHLSFROYVKGLAADGVCSPTPTULTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKKSRSSFGDSPSA VSPHLSFROYVKGLAADGVCSPTPTUTLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKKSRSSFGDSPSA VSPHLSFROYVKGLAADGVCSPTPTUTLGDA PVTNSSGKMALNSPQGPVESELGKQLLXTG WEGSPLPRSTPLOKAGATGAGVCPPASQGRPAGEP MGPEAGSKAELPPTVSRPLLARGLSWDSGFGE PGGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDPOLOPPASQGROPAGEP MGPEAGSKAELPPTVSRPLLAGLSWDSGFGE PGGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDPDISOVLLRKLREHILMRN QNLVGLKLPDLSEAAGPEKGLPSELSPALEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LITKEVENVFVQLSSAFRNOSYTLESSRINQAE RERNLTEENTEKLELPHKASITSSASLWHHCE		į	İ	1			
LITLITNTERRIPGEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHVKVRFKFYLLEFGYPAGTCFKDYVEING EKYCGERSQFVVTSNSNITUTRHSDQSYTDT GFLAEYLSYDSSDPCGFCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKL CKPLWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGBSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYDDRGFRYSPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHTPAGKAIWV TGWGHTQYGGTGALILQKGEHVNNQTTCEN LLPQQTFRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIPQAGVSMGGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSIMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNRPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFFQEFLAKEVSL EQGIVLPCRPPGGIPPAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHGGENVKGLALANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHGGENVKGLALANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHGGENVKGLALANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHGGENVKGLALANRVHSPHKRLSHRHLKVST WGGSPLPRSPTQDAAGVCPPASQGRGPAGEP MGPEAGSKAELPPTVSRPDLLRGLSWDSGPE PGPRLQKVLAKLPLLEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESKINQAE RERNLTEENTEKELENKASITSSASLWHHCE	}	ļ	}	j	ļ	•	· ·
RKAQGIFNSPYYLPGHYPNIDICTWNIEVPNN QHYKVRFKFYLLEPGYPAGTCPKDYVEING EKYCGERSQFVYTSNSNKITVRHSDQSYTDT GFLAEYLSYDSSDPCPQGFTCERGCIRKELR CDGWADACTDHISDELNGSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQIF RCSNGKCLSKSQQCADDCDGCDSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSTTRQARVVGGTD ADEGEWPWQVSLHALGGGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLIKRIISHFFNDFIFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLLQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGVDSCQGDSGGPL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGS SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTG EMTRPSLMAGRQHGWSAQQSATVANPVG ANPDLLPHIFLGEFEDVTYVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSGLP TMEVRINNSQQVEKVFGLEFYWCQCVAWS SSGTTKSQKAYKIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAG EQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTGGHSQQEAAMPHEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHTIDLYDPIGDISHSEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVTPTPFGDEADVS SPHFGEFNVFKGLADQLD,PIGISTEGENKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVTPTPFGDEADVS SPHFGEFNVFKGLADQLD,POSTSPEBCKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVTPTPFGDEADVS SPHFGEFNVFKGLADQLPGISTSGLKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVTPTSFGDEADVS SPHFGEFNVFKGLADQLPGISTPGGLAD PVTNSSGKMALNSPOPGPVESELGKQLLXTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPFTVSRPLLAEGLSWDSGFGE PGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAGQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRYHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENKASITSSASLWHHCE		İ		1			
OHYNVRPKEFYLLEPGYPAGTCPKDYVEING EXYCGERSQPVYTSNKITVRPHSDQSYTDT GELAEYLSYDSSDPCPGQFTCRTGRCIRKELR CCDGWADCTDHSDELNCSCDAGHQFTCKNKF CCPUPWCDSLNDGGDNSDEQGCSCPNAQTT RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK BDCSDGSDEKDCDGLSKFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYUDDGRFSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWY TGWGHTQYGGTGALLQKGEIRVNQTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVSCQGDSGGPL SSVEADGRIFQAGVSCQGDSGGPL SSVEADGRIFQAGVSCQGDSGGPL SSVEADGRIFQAGVSCQGDSGGPL SSVEADGRIFQAGVGSQQGATVANPVPG ANPDLLPHFLGEPEDVJVIVKNRVLLVCKAV PATQIFFKCNGEWQDVDVIERSTGGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTIKSQKAYIRAYLKNRFQEPLAKEVSL EQGIVLPCPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAEVPGTRGHSQQEAAMPHIPEDEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPIKRLSHRILKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHFGEPNVPKGLADRKQNDQRKVSQGRLAP RPPVVEKSKEIAJEQKENFPPLQYPETTPKGLA PVTNSSGMALLNSFRPOLYSPETTPKGLA PVTNSSGMALLNSFRPOLYSPETTPKGLA PVTNSSGMALLNSFRPOLYSPETTPKGLA PVTNSSGMALLNSFRPPLLRGLSWDSGPE MGPEAGSKAELPPTVSRSPPLLRGLSWDSGPE		Į.		[[,
### BEYOGERSOPVYTSNSNKITVRPHSDOSYTDT GFLAEYLSYDSSDPCPGGFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKE CCKPLFWVCDSLNDCGDNSDEGGCSCPAQTF RCSNGKCLSKSQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EBCSDGSDBKDCDCGLSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERKLRRIISPFFNDFTDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TOWGHTQYGGTGALLQKGEIRYNNQTTCEN LLPQQITPRMMCVGLSGGVDSCQDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWKENTGV ###################################	l	İ	1				1
GFLAEYLSVÖSSDPCPGQFTCKTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGICGASLISPNWL VSAAHCYUDDGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLKRIISHFFNDFTFDYDIALL ELEKPAEYSSMYRPICLPDASHVFPAGKAIWV TOWGHTQYGGTGALLQKGEIRYNNGTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVV TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRFSLMAGRQMWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVERINVSRQVGKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPFEGIPPAE 450 1800 A 3620 1 2676 MPSILGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQPTTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASTTSSRNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIALGERFPDPLQTPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELEPTVSRPPLLRGLSWDSGPEE POPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRNQKLTKLREEHILMRN QNLVGLKLPDLSEAAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRLLRVIRSLPGSAPP LTEKEVENVFVQLSSAFRNDSSYTLESRINVAE RERNLTEENTEKELELENFKASITSSASLWHHCE	1	!	ł	l l	-	ļ —	1 3
CDGWADCTDHSDELNCSCDAGHQFTCKNKE CKPLFWCDSLNDCGDNSDEGGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDGGRSTFRQARVGGTD ADEGGWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPIQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTDYDIALL ELEKPAEYSSMVRPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVNOTTCEN TGWGHTQYGGTGALILQKGEIRVNOTTCEN LLPQQITPRMMCVGFLSGVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSIAMAGRQHGWSAQQSATVAPVPG ANPDLLPHFLGEPEDVYIVNNKPVLLVCKAV PATQIFFKCNGEWVRQVOHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEYWCQCVAWS SSGTTKSQKAYIRAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSFTITVLTGDA TSPEGGTDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPA VSSNLSPSASPTSSRSNSLTVPTPFEGDEADVS SPHOEPNVPKGLADRKQNDQRRVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSSTQDAAGVGPPASQGRGPAGEP MGPFAGSKAELJRGLSWBSGFPE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREEHILMRN QNLVGILLPDLSEAAEQEKGLISPSLSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESSRINQAE RERNLTEENTEKELENKASITSSASLWHHCE	1	İ	1				
CKPLFWVCDSLNDCGDNSDEGGCSCPLAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLLKRISHPFFNDFTFDYDIALL ELEKPAEYSSWYPPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGFL SSVEADGRIFQAVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSFWRMDGSTERLEARRPAGRLFWSSRQ EMTREPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEBCHKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHOGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENPDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSFTQDAAGVGPASQGRGPAGEP MGPEAGSKAELLPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQQVQVVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAAQEEKGLPSELSPALEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESSRINQAE RERNLTEENTEKELENKASITSSASLWHHCE	1			1			
RCSNGKCLSKSQQCNGKDDGGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDGGLRSTTRQARVVGGTD ADEGEWPWQVSLHALGQGHIGGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERKLKRISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEFYWCQCVAWS SSGTTKSQKAYIRAYLRKNFEQEPLAKEVSL EQQIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTYPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEQEKGLPSELSPALEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENRKASITSSASLWHHCE	1		[[
KVNVYTCTKHITYRCLIGICLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDFIQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFITDYDIALL ELEKPAEYSSMYRPICLPDASHYPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQTTPRMMCVGFLSGGYDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSIMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTITKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGTLPCRPGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGPTTAGVSCSPTFITVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEDKKKN LALLEEAKLVSERFLTRGGRKSRSSPGDSPSA VSPNLSPSASFTSSRSNSLTVPTPPEGDEADVS SPHGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQFGPVESELGKQLLKTG WEGSPLPRSTTQDAAGVGPPASQGRGPAGEP MGPBAGSKAALPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVLAKLPLAEEEKRFAGKAGFAGEP MGPBAGSKAALPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKVPCVLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	ĺ	ĺ	!			CKPLFWVCDSLNDCGDNSDEQGCSCP\AQTF
### BDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERILKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHYPFAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV ### BT 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSILMAGRQHHWSAQQSATVANPVFG ANPDLLPHFLGEPEDVYIVKNPVLL VCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEK-VFGLEFYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFGEPLAKEVSL EQGIVLPCRPPEGIPPAE ### B 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGSQQEAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDNLAANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSFSASFTSSRSSNSLTVPTPPEGDEADVS SPHIPGEPNVPKGLADRKQNDQRK-VSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPPOGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGFRLQKVLAKLPLAEEEKRRAGKAGGKLAK APGLKDFQIQVOPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENNFVQLSAFRNDSYTLESRNQAE RERNLTEENTEKELEENFKASITSSASLWHHCE		1					RCSNGKCLSKSQQCNGKDDCGDGSDEASCP
ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAFGVQERLKRIISHPFFNDFITFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQVGGTGALILQKGEGVDSCQGDSGGPL LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNRPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPIPTITVLTGDA TSPEGETDKNLANRYHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHIGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSFTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APPGLKDFQQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLFGSAPP LTEKEVENVFVQLSSAFRRDSSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1	}]			KVNVVTCTKHTYRCLNGLCLSKGNPECDGK
VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGPLSGGVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVYY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFRCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTIKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVSTI ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHIGGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFFAGKAGGKLAK APGLKDFDQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLFGSAPP LTEKEVENVPVQLSSAFRRDDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	İ	!		i '			EDCSDGSDEKDCDCGLRSFTRQARVVGGTD
VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGPLSGGVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVYY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFRCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTIKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVSTI ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHIGGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFFAGKAGGKLAK APGLKDFDQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLFGSAPP LTEKEVENVPVQLSSAFRRDDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		į.	1				ADEGEWPWQVSLHALGQGHICGASLISPNWL
QRSAPGVQERLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMYRPICLPDASHVFPAGKAIWY TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVSI ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEJAIEQKENFDPLQYPETTPKGLA PVTNSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFFAGKAGGKLAK APGLKDOPQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		!	1				VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
ELEKPAEYSSMYRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQ/ITCEN LLPQQITPRMMCVGFLSGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTIKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVEENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	İ		1 1			
TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEY WCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSFTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGSPSA VSPNLSFSASFTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLIKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	ì	ŀ					
LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEVWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEKERFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPPLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	J	j					
SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TIMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVSI ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSFTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	ļ		ļ	1 1			
449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSAQQQPFTAGVSCSTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSFTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	!					
449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTINSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	ł	ł	ì		Í		
EMTRRPSLMAGRQHGWSAQQSATVANPVPĞ ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LITEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	140	1700	_	2619	2	412	
ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 A 3620 BEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLSEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	447	1799	^	3016	2	013	
PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLLSEEKFFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1		ļ				
TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	1	[
SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	1	l				
450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSFTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASFTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	i	l				
450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	}						
ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKONDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	150	1000		2600	L <u>. –</u>	0.00	
PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	450	1800	Α	3620	1	2676	
TSPEGÈTDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1 .	18	[
ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1					
LALLEEAKLVSERFLTRRĞRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	1					
VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	J					•	ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN
SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS
RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	1	ľ				SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP
PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	1					
WEGSPLPRSPTQDAAGVGPPASQGRĞPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE)				
MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1				ļ	
PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		!					
APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1		ļ			
QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		l	}				
ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	1	İ					
LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		ł					
RERNLTEENTEKELENFKASITSSASLWHHCE		1					
	Į į	i 1					,
HRETYQKLLEDIAVLHRLAARLSSRAEVVGA							
		L	L			·	HKETYQKLLEDIAVLHRLAARLSSRAEVVGA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	Α	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVQCQHLEALNEGTMG
453	1803	Α	3637	662	142	IQAKGLGIWHVPNKSPMQHWRKGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT \HVGYSSSREITE\AAVLLFYR
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457	1807	Α	3660		1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWIDEVHHGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQP\S LNQEDIYITTESLTNTAAGSP\GTGEHVPGSEM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS SCGYVSTDQLNKIMP
458	1808	Α	3663	154	462	TRAPASGRSGAGLALSANAPDSGGHPGATEG PAGSLAHASGSARGTWRVRGRGSHGWERTV GAGGCANPVPALHSCASAPRGTGRVSALGPK TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC \NQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR SA\NGNSGFQHETHAEETPNQPFNSVHLFSFM VLALNVVTVATITVRHFVNQRADYQ\YQKLQ NY
460	1810	Α	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\ TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV VQTGL*LLALSNPPALASQIAGISGMSHRAWP GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	A	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQTN CYYD/STKSFFYISCG*KVRKPTWAENRRLNA KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG HGS
463	1813	Α	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG GWRRJRPKSSTK
464	1814		3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT FSTPSSPAPFPTSSNPSSATTPPNPSP\GQR\DSR FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE AADGTRLDDQPKADVLEAHEAEAEEPEAGK SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR WHGEVAIRLLEMDGHNQDHLKLFKKEVMN YRQTRHENVVLFMGACMNPPHLAIITSFCKG RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF \GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR EMTPGKDEDQLPFSKAADVYAFGTVWYELQ ARDWPLKNQAAEASIWQIGSGEGMKRVLTS VSLGKEVSENLSACWAFDLQERPS\FSLLMD MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR FERFGLGVLESSNPKM
465	1815	Α	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP QAPLLAVGISFGGILVLNHLAQARQAAGLVA ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI
466	1816	A	3684	3	307	GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE GLPDLRALLPSEDRNS SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
467	1817		3687	2465	837	TGVLQG ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAQLVRLAGGHCRRDTLWKRLFLLE PPGPDRLRLGGRLALAELEELLEAVHAKSIGD IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY HHLESVINTACFTLWTRLL*GSGLDH*MSLFL ESWAYQIACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ TPPRAPLPESCPLVPLTTVSHLCPLSLRVFTSHL DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQQRN
468	1818	A	3691	960	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA SQNGGITGMSPCAWPFFFFFGLC
469	1819	A	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP HSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFILKYLLKTQENVYNIIEEVKKICSVLGCVE TKQITDAVNELSILILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSIPLVLGSAPGW DERTVSEMHTILRRWTFSQPLEALGILTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEPSKEQKLIKILGDIGERVKSASDHQRQEVL KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH DACSYFTSNALPLKITFINANLMGKNISIFKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					·	GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIJYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAFFIFTS EMEYFITEGGKNPQHFQDFVELCCRAYNIIR KHSQLLLINLLEMMLYAGLPELSGIQDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW
470	1820	Α	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVQLLLSLGSVFLWALL
471	1821	Α	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824	A	3753		. 5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSP\ERAALETPIQGQDGSPELLIRSLV GGPSAELLLDLER VLCREGSPGGAVRPLLKRL QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA	SEQ II		Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Source Seq	1		hod				
uence 1914 0.074906 0.0749001 0.0749001 0.074901 0.07			ł				
Berick Bot first amino acid of residue of misidue of peptide sequence S				1			
mino acid residue of peptide residue of peptide residue of peptide sequence feet		uence					
residue of popidie sequence Y=Tymsine, X=Unknown, **Siop codon, Y=Tymsine, X=Unknown, Y=Tymsin	uence			914			, , , ,
Popsible mucleotide deletion, Prossible nucleotide insertion nucleotide nucl		ľ	Ì	ĺ			,
			1			sequence	
ALRSGESGALLQOSFLTAMMSCOPARPINULSO QIGGGLIGGAPSEMJOQUAM QIGGGLIGGAPSEMJOQUAM LEQIGGLIGGAPSEMJOQUAM LEQIGGLIGGAPSEMJOQUAM LEQIGGLIGPARIS (POLMLOSLISTESLORGHALE QLQRLDKLFLQOEDEEKRI-FEEEEEEEE REKELFIEDPSPAISILVI.SPRCWPYSTLCITY PRICE		ŀ					
QIQGGLIGAFOYEM,GOLQRHLEPIMVLS LELATTFEHYOLMADRLISFSSWLEGAV LEQIGLCPPRILPQLMLQSLSTSEELGRQFUEL QLQRLDKLIFLGOEDEEKRLI FEBEFEEFEFEA RELFIEDPPAISILVI.SPRCVPVSPILCYLYH RKCLPTEFCDALDRYSSYSQSQNHPVLDMG PHRRLQWTWLGRAELQFGXGLILVSTYQMW LLLKFNQTEEVSVETLLKDSDLSPELLQAU, PLTSGOMEPLTHEGOPPHGGYLRLHEPOPQ RSGEALWLIPPQAYLAVERDEGTTLEQKIRL LCLLVRILKAHGEGLEHIQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCLHLLL GQGYVKRRDDRYDLMYADFPOPCRQA DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPOEVEGLMKQTVRQVOETLABPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLAAGLCV HQAQAVPVRPDICHYVSVEDGGTLEQKLALGYVE SCSNLTWCTNPQGCDRLLRQGLGCGTTCSK CGWASCPNCSFPEAHYPASCCHMSQWVDOD GYYDGMSVEAQSKHAKLISKRPSQAPIE KNEGCLHMTCAKCNHIGFCWCLLSWKRMH KDYYNCSAMVSKAARQEKRPQTVNERCTH HQAREFAVNLRNRVSAHEVPPPRSSTILDAD CQGLQARKVLAVACVSYSYQDAEYMDVV EQQTENLEHTNALQLLEETLLRCRDLASSL RLRADCLSTGMEKLARGYVE GQTENLEHTNALQLLEETLLRCRDLASSL RLRADCLSTGMELRGQBLAGGYGTGCSK RRADCHAVACVSSYSYQDAEYMDVV EQQTENLEHTNALQLLEETLLRCRDLASSL RLRADCLSTGMELRGQBLAGGYGDAEYMDVV EQQTENLEHTNALQLLEETLLRCRDLASSL RLRADCLSTGMELRGGRULLALGHSVAG FRVGLQSSVEAWEAKGFMMFGSQPAASSL RLRADCLSTGMELRGGRULLALGHSVAG FRVGLQSSVEAWEAKGFMMFGSQPAASSL RLRADCLSTGMELRGGRULLALGHSVAG FRVGLQSSVEAWEAKGFMMFGSQPAASSL SREPSLCVPSRTTGAFAGKAVLVSSATFUTT PSPRGSVEAWAGACGFMFGSQPASSE ATGRRRATRTQQRTAALLTDGTTKTGAAS SREPSLCVPSRTTGAFAGKAVLVSSATFUTT PFGRVFTAAGSSVEAWEAKGFMFGSQPASSC PROGRAFTSDVINILLRYPKYKLTNUSPAGPAARA RSGGGRWRNAFRGRWAGSGPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLIVPAP SRGQTLGVRTTGAASSLTVDRSSLWFGRSA RSGGGRWRNAFRGRWAGSGSSWEAGGWR PAQAPATSDVNINLLRXPKYKRTBLDLTYGFLGN TLWEEGRQRPPETLQPAR PAQAPATSDVNINLLRXPKYKRTBLDLTYGFLGN TLWEEGRQRPPETLQPAR PAQAPATSDVNINLLRXPKYKRTBLDLTYGFLGN TLWEEGRQRPPETLQPAR PAQAPATSDVNINLLRXPKYKRTBLDLTYGFLGN TLWEEGRQRPPETLQPAR PAQAPATSDVNINLLRXPKYKRTBLDLTYGFLGN FFGMOVSPCQAGVWHIDDSLQNLPPGFK RSYLSLPSSWDTWHVPPPRQANGCIPAN-RSP PYVIRATINSSHPLCRAGSSPWEAGGWSH APPOMDFTALLCFARKGCLPRQSEGGTLNLL PPFRANKPLVLUGACTRARLPSDLKCL TLWEEGRQRPPETLYPACSIPTYPCFMSC PPVNRSTKSPSPBMLTPTRWCSIPTYPCFMSC PPVNRSTKSPSPBMLTTPTRWCSIPTYPCFMSC PPVNRSTKSPSPBMLTTPTRWCSIPTYPCFMSC PPVNRSTKSPSPBMLTTPTRWCSIPTYPCFMSC PPVNRSTKSP	<u> </u>				sequence		
LELATTEHPYQHMADRILSFGSSWLEGD LEQIGLCPPNIPQLMIQSLSTSELQRQPHILE QLQRLDKLPLFQEDEEEKRI-FEEEEEEEEPRPHIP QLQRLDKLPLFQEDEEEKRI-FEEEEEEERPCYCLYLPH RKCLPTEPCDALDRFSSFYSQSQNHPVLDMG PHRRLQWTWLGRAELQFGKQILHVSTYQMU LLLKFNQTEEVSYSTLLXBDSLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPDQO RSGEALWLIPPQA/INVERDEGTTLEGKRNI. LSCLLVRILKAHGEKGI-HIDQLVCLVLEAW (KGPNPPGTLGHTVAGGVACTSTOVLSCILHLL) GQGYVKRRDDRPQILMYAAPEPMGPRCGQA DVPPCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGIMKQTYRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAVPVRPDICPVCVSPLGCDDLPSLCC MHYCCKSCWNEYLTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAQPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTRAAVSAAVACVSFYSQDAEPMGAAV EQQTENLELHTAALQULEETLEACCHASA CPAGPTGAFRAAVSTREQNILVINCTCHEA CPAGPTGAFRAAVSTRAAVACVSFYSQDAEPMGAAVACVSFYSCHAAVAC		į					
LEQIGLCPPRILPQLMLSLSTSEELORPHIA	Ì	i	1	ł			
Q.Q.R.D.K.I-I.G.Q.B.EEKR." SEEEEEEEEEEEEA EKEL.FIEDPSRISILLI.S.R.C.W.YSPI.C.Y. I.YH RKCL.PTECDAL.DRFSSFYSQSQNIFPYL.DMG PHRRL.Q.W.TW.L.G.R.EL.Q.FGK.Q.I.L.W.STV.Q.MW LLL.K.FNQTEEVS.V.ETL.L.K.D.S.D.S.P.ELLL.Q.A.L.V PLTSGNOPL.T.L.HE.G.Q.D.FT.D.L.C.P.G.R.S.G.E.A.W.L.P.P.Q.Y.L.N.W.B.G.T.L.B.C.L.V.R.L.K.AH.G.EK.G.H.ID.Q.V.C.L.L.E.A.Q. K.G.P.P.P.G.Y.L.D.F.P.Q.Y.L.N.W.G.R.T.L.B.C.L.W.R.C.L.B.C.L.W.R.C.L.B.C.L.B.C.L.B.C.L.B.C.L.B.C.L.B.Q.G.Y.W.R.D.D.R.P.C.R.G.C.L.B.C.L.		1.					l
				1			
RKCLPTECDALDRESSYSQSQNIPPULDMG PHRRLQWTWLORAELQFGKQILHVSTVQMW LLLKFNQTEEVSVETILKDSDLSPELLLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ RSGEALWLIPPQAYLNVENGETILEQKRIL LSCLLVRILKAHGERGLHIDQLVCLVLEAWQ KGPNPPGTIGHTVAGGVACTSTDVLSCHLHLL GQGYVKRRDDRPQILMYAAPEPMGPGTLQKRNIL LSCLLVRILKAHGERGLHIDQLVCLVLEAWQ KGPNPPGTIGHTVAGGVACTSTDVLSCHLHLL GQGYVKRRDDRPQILMYAAPEPMGPGTQA DVPPCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLAAGLCV HQAQAVPVTPHDCPVCVSBTCHDLPSLCDLDLSSLC MHYCCKSCWNEYLTTRIEQNLVLNCTCPJAD CPAQPTGAFRAVISSPEVSKYEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGGTTCSK GCWASCPNCSFPEAHYPASCHMSOWDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLIMTCAKCNHGFCWRCLKSWRNH KDYYNCSAMVSKAARQKRKPQDVNEECTTH HQAREFAVNLRNRVSAHEVPPRSSTTLNDA CQGLEQARKVLAVACVYSYSYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLIABGLTSGMELLRRIJCRSLALHJCHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPASSOP EAFEEEDDEDDVPEWQQDEFDEELDNDSFS VDESSNLDGETFFGDEEDEDEAYD 475 1825 A 3754 1093 96 GTSSNQISSKTHA*RSSWYPRPPCLIPLOPQ ATGRRRATITGQRTAALLTDGTTKTOAD SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPGSPTGAAGKLRAPGNRAGSPSSQEPPD DGTRRPASTTGQRTAALLTDGTTKTOAD SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNRSPFTAAGKLRAPGAK*AVLVRSATPTIN PPNRSPFTAAGKRAPATPSLPCLIPVAP SRGOTLOVRTTGASSLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHKCTYCTNHYSPAGFARS HVYIRRATINSHPLCRAGSPSSQEPPD DGTRRPASTTGVAGSPATRATPSLPCLIPVAP SRGOTLOVRTTGASSLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHKCTYCTNHYSPAGFARS HVYIRRATINSHPLCRAGSSPSSQEPPO GTRRPASTSDVNINLIRKPRVKRHDLTVQFLGN TLWEEGRQRPPETLQPAA 476 1827 A 3761 843 575 GVISAHCNLRLCHLGSSSNSPASASQVAGTIG ATTTPS*EYFLVETGFHHVSQDGLDLINFVI PRPRPLKVLGLQACTRARLSPLKEL 478 1828 A 3763 267 1240 HLISFILWSASIDCLEQLSGERIVKGNLLGP PPVNESTKPSPSWKLTPPMCSIPPYCFHYPSY WTGHSFASQVARLQYPESVKHLQCTSAESLL VINISHSSQUARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGHSFASQVARLQYPESVKHLQCTSAESL WTGH		-		Ì			
PHRRLQWTWLGRAELQFGKQLLHYSTVQMW LLLKFNQTEEVSVETLLKABSPELLQALV PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL LSCLLVYBLKAHGEKGHLIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSITOVLSCHLLL GQGYVKRDDRPQLMYAAPEPMGFCRGQA DVPPCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVGGLMKGVTKQVQCTLALEPDVAQH LLAISHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAYDYPDCHCVCYSPEAVATLASLQLPAGRT MSPQEVGGLMKGVTKQVQCTLALEPDVAQH LLAISHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAYDYPDCFVCVSPGCDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CPAQPTGAFRAIVSSPEVISKYEALLTGYVE SCSNLTWCTNPQGCDRLLCRQGLGCGTTCSK CGWASCFNCSFPEAHYPASCCHMSQWYDDG GYYDGMSVEAQSKHLAKLIKRCPSCQAPIE KNEGCLHMTCAKCHHGFCWRCLKSWRNH KDYYNCSAMVSKAARQKEAYQNSECTHHOW KDYYNCSAMVSKAARQKEAYQNSECTHHOW EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLGSPSVEAWEAKGFNGVNEECTHH HQAREFAVNLRNEVSAHEVPPPRSFTTLNDA CQGLEQARKVLAVACVYSYGOAEYMOVE EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLGSPSVEAWEAKGFNGSQPQASSGOP EAFEEEDDEDDVPFWQQDEFDEELDNDSFS YDESSENLDQEITFFGGEEEDOSGPQFASSGOP EAFEEEDDEDDVPFWQQDEFDEELDNDSFS YDESSENLDGEITFFGGEEEDOSGPPSQEPPD DGTRRPASTITGVAGASRATATPSLPCLHVVAP SRGGTLUFRTGARSTGVAGSPSTGPSSQEPPP DGTRRPASTITGVAGASRATVDSRSLSWPGRSA RSGGGRWRPAPARGRNFRAPFSSWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGFARS HVYIRATINSISHILCRAQSSAGTGVSHH PWRWFPFAAESPPHRCTYCTNHVSPAGFARS HVYIRATINSISHILCRAQSSAGTGVSHH VIRGHTRSSHFTCHCASSAGTGVSHH APPQMDFTTALLCFAPKGCLPRQKEGGTINLI FFFGNGVSPCPQAGV*WHDLDSLQNLEPGFK FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVCFTPM*RG FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVCFTPM*RG FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVCFTPM*RG FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVCFTPM*RG FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVCFTPM*RG FTMLARNVSIS*PRDLFALASQSAGTGSSHTHAPPSVFFTPSVCFTPHVSQDGLDLLIPFV PRPRLKVLQLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLSFRLWSASCDCLEQLSQGERVKGRALLGP PVVNESTKPSSPWKLTPPMCSIPVYPPRAKGLLGP FYTWSSPSGHSKLEVERAQTGPPCCHHYCP*P GVTDNTTSLLHYPPFCANCTIPPFSX WTGISFASQAALRQYPEVSKAGLLGP FYTWSSPSGHSKLEVERAQTGPPCCHHYCP*P GVTDNTTSLLHYPPFCANCTIPPFSX WTGISFASQAALRQYPESVKHLQCPSAESLL UTHERPOATERS WTGISFASQAALRQYPESVKH	1			1			
LILKPNQTEEVSVETILKDSDLSPHLIQALV PLTSGNORUTH HEGOPPRIGOVLRHEPOPO RSGEALWLIPPQAYLNVEKDEGRTLEQKRNI. LSCLLVRIKAHGEKGI.HIDQLVCVLIEAWQ KGPNPFOTIGHTVAGGVACTSTDVLSCILHLL GQGYVKRRDDRPQILMYAAPEPMGPCRQQA DVPFCGSGSSETSKPSEAVTALSLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAGQLLQSYSDFPLLLAAGLCV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWWSPLTTRIEQNLVLNCTCPIAD CPAQFTGAFRAVSSEPSVISYEKALLRGYVE SCSNLTWCTNPQGCDRLCRQGLGCTITCSK GCWASCFNCSSPEAHTPSAGGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMYSKAARQEKRFQDYNERCTFH HQAREFAVNLANIAVSAHEVPPPRSFTILMDA CQGLEQARKVLAYACVYSSYSQDAEYMDOV EQQTENLELHTNALQILLEETLLRCRLASSL RILRADCLSTOMELLRETULRCRLASSL RILRADCLSTOMELLRETULRCRLASSL RILRADCLSTOMELLRETULRCRLASSL RILRADCLSTOMELLRETULRCRLASSL RILRADCLSTOMELLRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMELRETULRCRLASSL RILRADCLSTOMER RICRATURY RILRADCLSTOMER RICRATURY	1	l		}			
PLTSGNGPLTLHEGQDEPHGGVLR,HEPOPQ RSGEALWIPPQAYLNEDGRTLEGVRNI LSCLLVRIKAHGEKGLHIDQLVCLVLEAWQ KGPNPGTIGHTVAGGVACTSTDVSCILHLL GGGYVKRRDDRPQILMYAAPEPMGPCRQQA DVPFCGSQSETSKPSFEAVATLASLQLPAGRT MSPGEVEGIMKQTVRQVOPTLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLAAGLCV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCKSCWMEYLTTREQNLVLNCTCPIAD CPAQTGAFIRAIVSSPEVISKYPEKALIRGYVE SCSNLTWCTNPQCDDRILCRQGLGCGTTCSK CGWASCFNCSPPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKSKRCPSQOAPD GYYDGMSVEAQSKHLAKSKRCPSQOAPD KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNBECTTH HQAREFAVALNRNVSAHISKRCPSCOAPD EQTTENLEHTNALQILISKRCPSCOAPD CGQTENLEHTNALQILISKRCPSCOAPD EQQTENLEHTNALQILISKRCPSCOAPD EQQTENLEHTNALQILISKRCPSCOAPD EQQTENLEHTNALQILISTRICTLACROLLASSL RILAADCLSTGMELLRRQGERLALDHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETTFFGDEEEDEDAYD GTSRNGHSPSTHAFASSWVQPPPLFLPLQPQ ATGRRRRRTRTQRAALLTDGTTKTQAAW SRRPSLCWPSSTTGAFGAVALVASATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITIOVAQSPATRATFSILCHIVPAP SRGQTLGVRTTGRASSRLTVDRSRLSWPGRSA RSGGGRWRNAPGRRWPRAP'SWPGRSA RSGGGRWRNAPGRRWPRAP'SWPGRSA RSGGGRWRNAPGRWPRAP'SWPGRSA RSGGGRWRNAPGRWPRAP'SWPGRSA RSGGGRWRNAPGRGWPRAP'SWPGRSA RSGGRWRNAPGRGWPRAP'SWPGRSA RSGGGRWRNAPGRGWPRAP'SWPGRSA RSGGRWRNAPGRGWPRAP'SWPGRSA RSGGRWRNAPGRGWPRAPTSCHLINGAPGRGWPRAPTS RTMLTATATA	1						, , , , , , , , , , , , , , , , , , , ,
RSGEALWLIPPQAYLNVEKDEGRIEQKRNIL LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL QGGYVKRRDDRPQILMYAAPEPMGPCRQQA DVPPCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLAILEDDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAVPVRPDICPVCVSPLGCDDDLPSLCC MHYCKSCWMEYLTTBGUN,ULNCTCPIAD CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE SCSNLTWCTNPQGCDRLCRQGLGCGTTCSK CCWASCFORSPEAHTYPASGGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYVNCSAMVSKAARQEKRFQDYNERCTHH HQAREFAVALRAWVSAHVPPRSFTEINDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELLRETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RLLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALQILLEETLIKRCBLASSL RTLRADCLSTOMELHTNALGILLETLIKRCBLASSL RTLRADCLSTOMELHTNALGILLETLIKRCBLASSL RTLRADCLSTOMELHTNALGELLAILETLIKRCBLASSL RTLRADCLSTOMELHTNALGILLETLIKRCBLASSL RTLRADCLSTOMELHTNALGLASSL RTLRADCLSTOMELHTNALGELLAILETLIKRCBLASSL RTLRADCLSTOMELTATIONELTH RTTLTTS-IFVFLVETGHPHVSQPGLONELTHNALGELIKRCBLITATIONELTHTPMCSIPPVFPRSGS PTSSWASPSGMSKLEVERAQTGPFCLHITYCPP GYTDNTTSLHTYBFPRLSGLCVCFPAH*FFSY WTGHSFASQAWLRQVPSSESLL VTGHSTAMELTER RTLRADC	1			1			
LSCLLVRILKAHĞEKGLHIDQLVCLVLEAWQ KĞPNPPGTIGHTVAGVATSTDVLSCILHLI GQGYVKRDDRPQILMYAAPEPMGPCRGQA DVPFCGSQŞETSKPSPEAVALLASLQLPAGRT MSPQEVEGILMÇIVYRQVQETINLEPDVAQH LLAHSHWGAEQILQSYSEDPEPLLLAAGI.CV HQAQAVPRPDHCYSPLGVDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CRAQPTGAFRAIVSSPEVISK YEKALLRGYVE SCSNLTWCTNPQQCDRILCRQGLGGGTICSK CGWASCFNCSFPEAHYPASCGIMSQWVDDG GYYDGMSVEAQSKH.KLISKCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKFQDYNERCTIFH HQAREPAVNLRNRVSAHEVPPPRSFTIFLNDA CQGLEQARKVLAYACVYSPYSQDAEYMDV EQQTEBLEHTINALQILLETTLIRCRDLASSL RILIRADCLSTGMELLRIQERLLAILQHSAQD FRVGLQSPSVEAWEACPNMPGSQPOGASSGP EAEEEEDDEDDVPEWQQDEFDEELDNDSFS YDBSENLQGTFFFGDEEDBEDEAYD FRVGLOSPSVEAWEACPNMPGSQPOGASSGP EAEEEEDDEDDVPEWQQDEFDEELDNDSFS YDBSENLQGTFFFGDEEDBEDEAYD ATGRRRRRTRTQQRTAALLTDGTTKTGAAW SRPSLCWPSRTTGAKA*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGRRAGSEPSSQEPPP DGTRIRPASTIGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASKLTVDRSRLSWPGRSA RSGGGRWRPMAPRGRWPRAP'SWEPGSWTE PWWPPFPALSSPHRCTYCTNHVSPAGPARS HVYIIRATINISISHPLCRAQSSPWEAAGVWRR PAQPAFTSDVNINLLRYRVKRHDLIYQFLGN TLWEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGRGVSPCPQAGV*WHDLDSLQNLEPGFG FTMLARMVSIS*PRDLPALASQSAGITGYSHH APPQMDFTFALLCTPAGCLPRQKEGGTINLI 477 1827 A 3761 843 575 GVISAHCNIRL/CHLPGSSNSPASASQVAGTIG ARTTS*FIPVILVETGFHHYSQDGLDLLNPVI PRRFILKVLGIQACTRARLFSPIKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTREPSSPWSTLPPRCSISPPPPKSGS PITSWSPSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHVPFPRUSGLVCPPAHFPSS WTGHSTRASQAWLRQVFVSKHLCCPPAHFPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS WTGHSTRASQAWLRQVFVSKHLCVCPPAH*PPSS	İ		ŀ	İ			
KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL GQGVVKRRDDRPQLMYAAPEPM6PCRQA DVPFCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETTLNEPDVAQH LLAHSHWGAGQLLQSYSEDPEPLLLAGLCV HQAQAVPVRPDHCPVCVSPLGCDDLPSLCC MHYCCKSCWMEVITLEQMI VINCTCPHAD CPAQPTGAFIRAIVSSPEVISK YEKALRGYVE SCSNLTWCTNPQGCDRLLCRQGLGCGTTCSK CGWASCPNCSFPEAHPYASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHFGWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPRSFTFLNDA CQLEGQARKVLATVACVYSPYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLAGGELLALIQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDEELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDEELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDEELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDEELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEDDEDDVPEWQQDEFDEELDNDSSS FRVGLQSPSVEAWEAKGPNMPGSQPQASSOP FRVGLQSPSVEAWEAKGPNMPGSCPPCFPEPLPLOPQ ATORRRETTRTQQRTAALTDGTTKTGAAW SRPSSLCWPSRTTGAGAK*AVLVNSATPTIN PPNPQSPTGAAGKLARPGNRAGSEPSSQEPS PWRWPFPAAESSPPHRCTYCTTNHVSPAGPARPS HYVIIRATINISISHLCRAQSSPWEAAGVWRR PWRWPFPAAESPPHRCTYCTTNHVSPAGPARPS HYVIIRATINISISHLCRAQSSPWEAAGVWRR PWRWPFPAAESPPHRCTYCTTNHVSPAGPARPS HYVIIRATINISISHLCRAQSSPWEAAGVWRR FFFGNGVSPCPQAGV*WHDIDDSLQNLPPGFK RSYLSLPSSWDYRHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHDIDDSLQNLPPGFK FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQANFCIFM*RG FFFGNGVSPCPQAGV*WPHPRQA	ł]		ŀ			
GQGYVKRRDDRPQILMYAAPEPMGPCRGQA DVPPCGSQSETSKPSPEAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV HQAQAYPAPPHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD CRAQPTGAFRAIVSSPEVISK YEKALLRGYVE SCSNLTWCTNPQQCDRLCRQGLGGGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHALLSKRCPSCQAFIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSFTEINDA CQGLEQARKVLAYACVVSFYSQDAEYMDVV EQQTENLEHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRQERLLALQHSAQD FRVGLQSFSVEAWEAKGPMMPGSQPQASSGP EAEEEEEDDEDDVPSQDEFDEELDNDSFS YDESENLDQSTFFFGDEEEDEDAYD ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPMQSPTGAAGKLRAGPRAGYGSPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVVBGRASSHVEAGSAPARS ASGGGRWPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTVCTNHVSPAGPARPS HYVIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPATTSDVNINLLRICRVKRHDLTYQFLGN TLWEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGRGVSSPCPAGYWHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RG FTMLARMVSIS*PRDLPALASQSAGTIGVSHTM APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPGMDFTALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKFCAPKGCLPGCHTQCFP GVTDNTTSULHYVPPPRLSIDLVCFPPPPPKSGS PITSWSPSGHSKLEVERAGTG	Į						
DVPPCGSQSETSKPPBAVATLASLQLPAGRT MSPQEVEGLMKQTVRQVQETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAGGLCV HQAQAVPVRPDHCPVCVSPLGODDLPSLCC MPYCCKSCWNEVLTTRIEQNLVINCTCPIAD CPAQPTGAFRAIVSSPEVISK YEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCRNCSFPSAHYPASCGIMSQWVDDG GYYDGMSVEAQSKHLAKLISKCPSCQAPIE KNEGCLMTCAKCNHGFCWRCLKSWEPNH KDYYNCSAMVSKAARQEKRFODYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSFTFLNDA CQGLEQARKVLAVACVVSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RILRADCLSTGMELLAGPELLALLQHSAQD FRYGLQSPSVEAWEAKCPNMPGSQPOASSGP EAEEEEDDDVPEWQODEFDEELDNDSFS YDESENLDQETFFFGDEEDEDLAYD FRYGLQSPSVEAWEAKCPNMPGSQPOASSGP EAEEEEDDVPEWQODEFDEELDNDSFS YDESENLDQETFFFGDEEDEDLAYD ATGRRRRTRTQRALALTDGTTKTGAAW SRRPSLCWFSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAGSEPSSQPPP DGTRIRAFRSTGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGASRLTVDRSRLSWPGRSA RSGGG MYPNPAPRGSQFTAATPSLPCLHVPAP SRGQTLGVRTTGASRLTVDRSRLSWPGRSA RSGGG MYPNPAPRGSRTSPSSQPSWEAAGVWRR HVYIIRATINSISHLCRAQSSPWAAGVWRR HVYIIRATINSISHLCRAQSSPWAAGVWRR PVWRPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHLCRAQSSPWAAGVWRR PVRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHLCRAQSSPWAAGVWRR PROPAPTSDVNINLLRKPRVKRHDLTYQFLGN TLWEGRQRPPETLQPAR PSYGGNSPSCPQAGVWHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGTIGVSHH APPQMDFTFALLCFAGCLPRQKEGGTINLI APPQMDFTFALLCFAGCLPRQKEGGTINLI APPQMDFTFALLCFAGCLPRQKEGGTINLI APPQMDFTFALLCFAGCLPRQKEGGTINLI APPQMDFTFALLCFHLPGSSNSPAASQVAGTIG ARTTPS*FVFLVETGFHHVSQDSLDLLNFVI RRRPLKVLJQLQCTRARLPSPLKEL A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGF PVNESTKPSPSWLACVCPRAHPPSY GVTDNTTSLLHVTPFRUKSICVCPPAPPSGS PTTSWSPSGHKKLEVERAQTGFFCLHIYCP*P GVTDNTTSLLHVTPFRUKSICVCPPAPPSGS PTTSWSPSGHKKLEVERAQTGFFCLHIYCP*P GVTDNTTSLLHVTPFRUKSICVCPPAPPSGS PTTSWSPSGHKKLEVERAQTGFFCLHIYCP*P GVTDNTTSLLHVTPFRUKSICVCPPAPSGSSLL WTCHSFASQAWLRQVFEVSKHLQCYSASSLL WTCHSFASQAWLRQVFEVSKHLQCYSASSLL WTCHSFASQAWLRQVFEVSKHLQCYSASSLL WTCHSFASQAWLRQVFEVSKHLQCYSASSLL WTCHSFASQAWLRQVFEVSKHLQCYSASSL	ŀ		İ				
MSPOEVEĞLMKQTVRQVÖETLNLEPDVAQH LLAHSHWGAEQLLQSYSEDPEPLLLAGGLCV HQAQAVPNEPDHEVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLITRIEDNLVINCTCIPLD CPAQPTGAFIRANISSPEVISKYEKALIRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDOMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSFTEINDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENILEHTHALQILLEFILLRCRDLASSL RLIRADCLSTGMELLRRIQERLLAILQHSAQD FRYGLQSPSVEAWEAKGPNMPGSQPQASSOP EAEEEEEDDEDDVFEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDEAYD GTSRNQHSPETHAFARSSWPQPPPFLEPPLOPQ ATGRRRRITTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKAVALVRSATPTIN PPNPQSPTGAAGKAVALVRSATPTIN PPNPQSPTGAAGKARAFGRAWPRAP'SWEPGSWTE PWRWFFPAAESPPHRCIYCTNHVSPAGPARS RSGGGRWRPNAFGSWPAAP'SWEPGSWTE PWRWFFPAAESPPHRCIYCTNHVSPAGPARS HVYIIRATNISSHPRICIPCAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLONLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFAM*RG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCTARKSCUPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNIRLICHLPGSSNSPASASQVAGTTG ARTTPS*HVYLUFGFHHVSQDGLDLINFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEGLSGERHVKGMLIGP PPVNESTKPSSPSWKLTPPMCSIPPPPPKSGS PTTSWSPSGHSKLEVERAQTGFFCLHIYCP*P GYTDNTTSLLHYIPFPRLISGLVCFPAH*FPSY WTGHSFASQAWLRQVEVSKHLQCPSAESLL			Ì	}			
LLAHŠHWGAEQLŪ, GSVŠEDPEPLLI, AAGLČV HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWMEYLTIRIEQNI, VI, NCT. CPIAD CPAQPTGAFRAIVSSPEVISK YEKALLRGYVE SCSNLTWCTNPGCDRILCRGGLGGGTTCSK CGWASCFNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGGLHMTCAKCNHGFGCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSTFILNDA CQGLQARKVLAKCVYSFYSQDAEYMDVV EQQTENI, EHTNALQILLESTLIRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDBEDDVPEWQQDEFDEELDNDSFS YDSSENLD, QETFFGDEEDEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRITRTQQRTAALLTDGTTTKTGAAW SRPSSLCWPSRTTGAPGAK*AVLVNSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASTITGAPGAK*AVLVNSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASTITGAPGAK*AVLVNSATPTTN PWRWPFPAAESPPHRCTYCTNHVSPAGPARS RSGGGRWRNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARS RYSGGTAWRNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARS RYSYLSLPSSWDYRHIPPRQANFCIFAM*RG FTMLARMVSIS*PLDALAASQSAGTIGSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHIPPRQANFCIFAM*RG FTMLARMVSIS*PLDALAASQSAGTIGSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRIJCHLPGSSNSPASASQVAGTIG ARTTPS*IFVLVBTGFHHVSQDGLDLL/NFVI RPRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLISHLUKASSLDCLEQLSGRHVKGMLIGP PPVNESTKPSSPSWLTPPRWSGRPVFPPRSGS PTTSWSPSGHSKLEVERAQTGPFCLHIYCP*P GVTDNITTSLLHYIPFPRLISGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL WTGHSFASQAWLRQVPEVSKHLQCPSAESLE	1			[
HQAQAVPYRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTRIEQNILVLNCTCPIAD CPAQPTGAFRANVSSPEVISK YEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGGGTTCSK CGWASCFNCSPEPAHYPASCGHMSQWYDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKARQEKRRQDYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQTENLEHTINALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQFRLLAILQHSAQD FRVGLQSFSVEAWEAKGPNMPGSQPQASSGP FAEEEEDDEDDDYPWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDDEAYD GTSRNQHSPKTHAFASSWPQPPPFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKAKAPAVRASSPESSQEPPP DGTRRPASTIGVAQSPATRATPSLPCLHVPAP SRGGTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWFFFAAESPPHRCTYCTNHVSPAGPARSF HVYIIIRATTNISSHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQAG 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPSASQVAGTTG ARTTPS-TFYLLVEGSFHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEGLSGRHVKGMLLGP PPVNESTKPSSPSWLTPPMCSIPPYPPPKSGS PTTSWSPSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRLSGLVCFPAH*FPSY WTGHSFASQAWLRQVFEVSKHLQCPSAESLL	Į.						
MHYCKSCWNEYLTTRIEQNI.VLNCTCPIAD	}	ļ		}			
CPAQPTGAFRAJVSSPEVISKYEKALLRGYVE SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK CGWASCPNCSFPEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAHEVPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLIRADCLSTGMELLRHQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFGDEEDEDEAYD GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAYGVRAAGSEPSSQEPPP DGTRRPASTIGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGASRLTVDRSRLSWPGRSA RSGGRWRPNAPRGWPRAP*SWPGGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRNVPRQAPTSDVNINLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRNVPRQAPTSDVNINLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 477 1827 A 3761 843 575 GYISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVLVETGFHHYSQDGLDLLJNFV RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVMESTKPSPSFWKLTPPMCSIPPVFPRKSGS PTTSWS/PSGHSKLEVERAQTGGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESSLL WTGHSFASQAWLRQVPEV							
CGWASCFNCSPFEAHYPASCGHMSQWVDDG GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTINALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLALIQHSAQD FRVGLQSSSYEAWEAKGFNNPGSQPQASSGP EAEEEEDDEDDVPFWQQDEFDEEDDNDSFS YDESENLDQETFFFGDEEDDEAYD GTSRNQHSPSYHAWEAKGFNNPGSQPQASSGP EAEEEEDDEDDVPFWQQDFFDEEDDAYD GTSRNQHSPSYHAWEAKGFNNPGSQPASSGPP DGTRRPASTTGVAQSPATRATPSLPCLHVPAP SRGGTLGVRTTGRAAGKLAPGGNRAG/SEPSSQEPPP DGTRRPASTTGVAQSPATRATPSLPCLHVPAP SRGGTLGVRTTGRAAGKLAPGGNRAG/SEPSSQEPPP DGTRRPASTTGVAQSPATRATPSLPCLHVPAP SRGGTLGVRTTGRAAGKLAPGGNRAG/SEPSSQEPPP HYYIIRATTINSISHPLCRAQSSPWEAAGVVRR PAQPAFTSDVINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGVVWHDLDSLQNLPPGFK RFSYLSLPSSWDYRRIVPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGTTGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI GVISAHCNLRL/CHLPGSSNSPASASQVAGTTG ARTTPS*IFVFLVETGFHHVSQDGLDLLNFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVPPPKSGS PTTSWJPSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFFRLISGLVCFPAH*FPSY WTGHSFASQAWLRQVFEVSKHLQCPSAESLL							CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE
GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE KNEGCLHMTCAKCNHGFCWRCLKSWKPHH KNDYNNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENILLEFTLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSFSVEA WEAKGPNMPGSQPQASSGP EABEEEDDEDVPEWQDEFDEELDNDSFS YDESENLDQETFFFGDEEDDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSSWPQPPFLF.PPL.QPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITTOFARSRLTVDRSRLSSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHCTYCTNHVSPAGPARPS HYVIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSIPSSWDYRYPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GYISAHCNLRL/CHLPGSSNSPASASQVAGTTG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HILSFHLWSASLDCLEGU_SQERVKGMLLGP PVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGFFCLHIYCP*P GYTDNTTSLLHYIPPFRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	ļ		1				SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
KNEGCLHMTCAKCNHGFCWRCLKSWKPNH KDYYNCSAMVSKARQEKRFQDYNERCTFH HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTINALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLALLQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESSINLDQETFFFGDEEDEEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRITTQQRTAALLITDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLHVPAP SRGGGRWRNAPRGRWPRAP*SWEFGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFARKGLLGP FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFARKGCLPRQKEGGTLINLI APPTS*IFVFLVETGFHHVSQDGLDLL/NFVI RFRRFLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERTWKGMLLGP PVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGGFFCLHIYCP*P GVTDNTTSLLHYIPFPRL/SGLVCFPAH*FPSY WTGHSFASQAWLRQVYEVSKHLQCPSAESLL	ļ						
KDYYNCSAMVSKAARQEKRFQDYNERCTFH HQAREFAVNLRINRYSAHEVPPRSFITELNDA CQGLEQARVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETLLRCRDLASSL RLLRADCLSTGMELLRRIQERLLAHLQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEDEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFPPLQPQ ATGRRRRITRTQQRTAALLITDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRIRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HYVIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPAALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PITSWS/PSGHSKLEVERAQTGFFCLHIYCP*P GVTDNTTSLLHYIPPFRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
HQAREFAVNLRNRVSAİHEVPPPRSFTFLNDA CQGLEQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTINALOİLLEETLLRCRDLASSL RILRADCLSTIGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRITRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTIGAPGAK*AVLVRSATPTIN PPNPQSYFGAAGKLRAPGARAG/SEPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRITTGRASRLTVDRSRLSWPGRSA RSGGGRWPPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLINLI APPQMDFTFALLCFAPKGCLPRQKEGGTLINLI APPQMDFTFALLCFAPKGCLPRQKEGGTLINLI APPQMDFTFALLCFAPKGCLPRQKEGGTLINLI APPROMDFTFALLCFAPKGCLPRQKEGGTLINLI APPRIFY-IFVFLVETGFHHVSQDGLDLL/NFVI RPRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWSPSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYPPFPLISGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESIL			1				KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
CQGLQARKVLAYACVYSFYSQDAEYMDVV EQQTENLELHTNALQILLEETILRCRDLASSL RLLRADCLSTGMELLRRIQERILAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRITIGAPGAK*AVLVRSATPTIN PPNPSQPTGAAGKLRAPGMRAG/SEPSSQEPPP DGTR\(\text{RPASITGVAQSPATRATPSLPCLHVPAP}\) SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLR\(\text{RRPRVRHDLYQFLGN}\) TL\(\text{WEGRQRPPETLQPAR}\) FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPGMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQKEGGTLNLI APPCMDFTFALLCFAPKGCLPRQCFCCHTYCVI RPRRPLKVLGLQCTRARLPSPLKEL APPCMDFTFALLCFAPKGCLPRQCGMCMGMLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHTYCP*P GVTDNTTSLLHYPPFRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVFEVSKHLQCPSAESLL							KDYYNCSAMVSKAARQEKRFQDYNERCTFH
### EQQTENLELHTNALQILLEETLLRCRDLASSL RILRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFGDEEDEDEAYD ###################################							HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
RLLRADCLSTGMELLRRIQERLLAILQHSAQD FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRITRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRIRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW/DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPAASSQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVPPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRLISGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL]		,		*		CQGLEQARKVLAYACVYSFYSQDAEYMDVV
FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS YDESENLDQETFFFGDEEEDEDEAYD 475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSSWPQPPPLFLPPLQPQ ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASTIGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HYYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI FFFGNGVSPSIFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRLISGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	İ						
### FEREEEEDDEDDVPEWQQDEFDEELDNDSFS #### YDESENLDQETFFFGDEEDEDEAYD ###################################	1	İ					
YDESENLDQETFFFGDEEDEDEAYD	1						`
475 1825 A 3754 1093 96 GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ ATGRRRRITRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRIRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSSAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRLSGLVCFPAH*FFSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL		1					
ATGRRRRTRTQQRTAALLTDGTTKTGAAW SRRPSLCWPSRTTGAPGAK*AVLVRSATPTIN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTRRPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHYSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	176	1006		2764	1000	06	
SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTR/RRASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAG*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI FFRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	4/3	1825	A	3/34	1093	96	
PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP DGTR\RPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HYYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS\PSGBHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
DGTR\RPASITGVAQSPATRATPSLPCLH\PPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCTYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
RSGGRWRPNAPRGRWPRAP*SWEPGSWTE PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI RPPRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							3
HVYIIRATINSISHPLCRAQSSPWEAAGVWRR PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTFALCFAPKGCLPRQKEGGTLNLI APPQMDFTALCFAPKGCLPRQKEGGTLNLI APPQMDFTALCFAPKGCLPRQKEGGTLNLI APPQMDFTALCFAPKGCLPRQKEGGTLNLI APPQMDFTALCFAPKGCLPRQKEGGTLNLI APPQMDFTALCFAPKGCLPRQCAGGTCAGLPRQCAGGTLNLI APPQMDFTALCFAPKGCLPRQCAGGTCAGLPRQCAGGTCAGLPRQCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGGTCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGAGCAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG							
PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSWDYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	l	1					
TLWEEGRQRPPETLQPAR 476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFV1 RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
476 1826 A 3758 901 521 FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFV1 RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
RFSYLSLPSSWDYRHVPPRQANFCIFM*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL ARROWSIS*PROBLES ARROWSIS*PROBLES ARROWSIS*PROBLES ARROWSIS*PROBLES ARROWSIS*PROBLES BYTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	476	1826	Α	3758	901	521	
FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFV1 RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI 477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PVNESTKPSPSPWKLTPPMCSIPPVFPKSGS PITSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
477 1827 A 3761 843 575 GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	L]					
ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	477	1827	A	3761	843	575	
RPRRPLKVLGLQACTRARLPSPLKEL 478 1828 A 3763 267 1240 HLLSFHLWSASLDCLEQLSQERHVKGMLLGP PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	}]		
PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	<u></u>						RPRRPLKVLGLQACTRARLPSPLKEL
PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	478	1828	Α	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY WTGHSFASQAWLRQVPEVSKHLQCPSAESLL							
WTGHSFASQAWLRQVPEVSKHLQCPSAESLL	[ĺ		ĺ	PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P
			·		1 1		
	l				!		
		[]			ļ		-
PDEHPQDTDARDADGEAREREP/RRPSFAA*P	l					ļ	
VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA	l				ľ	i	
CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE	1		1		,		
QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA	L	l					QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRFILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF I/ANFIMAKSWKQPTHPSVRT
481	1831	A _.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	Α	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	<i>.</i>	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP A*FFVFI.VE\QGFTMLARMVSIS*PQ/CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	Α	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	Α	3811	378	98	RYD*SSQSENIPAQKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLLYHLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771		FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FRACLLELIPYAPTLSWTACPPAMAGPRGLLP LCLLAFCLAGFSFVRGQVLFKGCDVKTTFVT HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR YEVQLGGSMVSMSGCRRKCRKQVVQKACCP GYWGSRCHECPGGAETPCNGHGTCLDGMDR NGTCVCQENFRGSACQECQPPNRFGPDCQSV CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
400	1839		2822	024	((0)	QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1039	Α	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG QDGLDLLNLMIHPPRPPKVLGFQA
490	1840	A	3825	79	9748	GCQSCWPAWPRLRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPPPQLPQPPPPQAPLLPQPPPPPPPPPP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	location corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		•		peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		YTTTCVQLQKQVLDLLAQLVQLRVNYCLL
						DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
	· ·					KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
						LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
						CHKENEDKWKRLSRQIADIILPMLAKQQMHI
						DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED
						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
						EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
						KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
						GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
						ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE
						HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
						AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
						HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
		·				ACRRVEMLLAANLQSSMAQLPMEELNRIQEY
İ						LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
						SQCWTRSDSALLEGAELVNRIPAEDMNAFM
						MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
}						AREVTLARVSGTVQQLPAVHHVFQPELPAEP
Ì	:					AAYWSKLNDLFGDAALYQSLPTLARALAQY
	•					LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
l i						SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
						EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
						TAACEMVAEMVESLQSVLALGHKRNSGVPA
						FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
						WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME
						QEESPPEEDTERTQINVLAVQAITSLVLSAMT
						VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
				:		LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
						PVPSLSPATTGALISHEKLLLQINPERELGSMS
					•	YKLGQVSIHSVWLGNSITPLREEEWDEEEEE
						ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS
						DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
			J			YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
]]		,				RSSHLPSRVGALHGVLYVLECDLLDDTAKQL
				ļ		IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA
		.			İ	TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEQLSRLDAESL
						VKLSVDRVNVHSPHRAMAALGLMLTCMYT
[[[ſ	ĺ	ĺ	GKEKVSPGRTSDPNPAAPDSESVIVAMERVS
						VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ
		. [DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT
					1	LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA TWSLSCFFVSASTSPWVAAILPHVISRMGKLE
		}			i	QVDVNLFCLVATDFYRHQIEEELDRRAFQSV
		1				LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
100		<u>.</u>	-			HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\FF FQSEWTAVV/P/EFTATQSEVADWFKDMQVP
}		ĺ			ĺ	SVPIQQFPTEDWST*PTMNDWSATSTAQTTE
<u> </u>			ŀ			WVRITTEWP
ــــــــــــــــــــــــــــــــــــــ					t	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq- uence	ł	09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	4.0		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		[l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ		İ	peptide sequence	ł	/=possible nucleotide deletion, \=possible nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
				} ~~		KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK
			}	1		CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
494	1844	A	3845		250	VCHLLAIKLGFYIEIHLTTFNNTF
494	1044	Α.	3643	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
						FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
						ARPQDIDFLYAHQGRCWFRLL
495	1845	Α	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
					ļ	WADKYRPRKPRFFNRVHTGFEWNKYNQTHY
						DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL EACADNKDFAILRFHAGPPYEDIAFKIVNREW
						EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
						RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
						QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
						HGELRRHWDRLA*GPDATEGALGASFEHEG GQQPPADLTVQADTLHRPSARLGGAHRACPK
						RRPHRVLWRWARGAWAWRCOAREKOETOG
						QPCHITGHPLGREAEPAAAGAAPALAHRPPF
						ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
						WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
						VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
1						TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
						LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
	1					SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
496	1846	A	3849	830	442	SFVLMELAYWQDRMFF AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
170	1040	A	3049	030	442	LLSSWDYRSLPPRPVNFCILVELGFHHVDOAG
			·			LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
						RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	Α	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
,						LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/POLLFVPR
[ſ	-		RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
li						VPIL
498	1848	Α	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG
						EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF*FLVETGFHHVGQADLKLLTSGDP
1						PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	A	3863	423	263	·APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
		.				KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
						KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
	-		1		1	MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
						LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP
			1			DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
1		[ĺ	VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
	ĺ		Ì			TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSOGGGHRTLLYGHAILLRHAH
]		ļ		ļ	SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
			j			ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
			Ì			SERYLHLSTASGELQVDASFMQTLWNMNPIC
						SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
	1					DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG
ļ	ļ	j				LVVVDASKAHTKATSFCFRISKEKLDVAPKR
			l			DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ľ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
ucite			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
İ						APDPKALRLGVLKKKAMLHQEGHMDDALSL
						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
						GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
						EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
						VNLLYELLASLIRGNRSNCALFSTNLDWLVS
						KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
						RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
1						ATHLRVGWALTEGYTPYPGAGEGWGGNGV GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
						LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
						NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
						LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
						LERIREKLAENIHELWALTRIEQGWTYGPVRD
			i i			DNKRLHPCLVDFHSLPEPERNYNLQMSGETL KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
						NGHNVWARDRVGQGWSYSAVQDIPARRNPR
						LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
						QSGRWYFEFEAVTTGEMRVGWARPELRPDV ELGADELAYVFNGHRGQRWHLGSEPFGRPW
	' i					QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
						ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
						VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
	' I			ĺ		KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
						LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
1			İ			NLRRSAGGWSEAENGKEGTAKEGAPGGTPO
1				ļ	ł	AGGEAQPARAENEKDATTEKNKKRGFLFKA
l i		1	İ		{	KKVAMMTQPPATPTLPRLPHDVVPADNRDD
		1				PEILLNTTTYYYSVRVFAGQEPSCVWAGWVT
		ľ				PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
						VIGCLVDLATGLMTFTANGKESNTFFQVEPN
			1			TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
			1		[AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
			I			MPNHFLQVETRRAGERLGWAVQCQEPLTMM ALHIPEENRCMDILELSERLDLQRFHSHTLRL
			l			YRAVCALGNNRVAHALCSHVDQAQLLHALE
	ļ					DAHLPGPLRAGYYDLLISIHLESACRSRRSML
	1	Ì	l			SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
1		- 1				GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
		Į				ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
	}	l	1			RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDEE
	İ		- 1		}	EKEEDEETAQEKEDEEKEEEAAEGEKEEG
]	j	İ	ļ			LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
	l	ł	İ			HRVESLAAFAERYVDKLQANQRSRYGLLIKA
						FSMTAAETARRTREFRSPPQEQINMLLQFKDG
			İ			TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD GEEEEPEETTLGSRLMSLLEKVRLVKKKEEK
		ł	-	1	1	PEEERSAEESKPRSLQELVSHMVVRWAQEDF
					1	VQSPELVRAMFSLLHRQYDGLGELLRALPRA
		l			l	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
1 1	i	1	1	l	!	ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
						TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l i				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						CYFCRISRQNQRSMFDHLSYLLENSGIGLGM
						QGSTPLDVAAASVIDNNELALALQEQDLEKV
						VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
1	i					ERYLDFLRFAVFVNGESVEENANVVVRLLIR
						KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
!						DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
İ						FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
1						LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
1 1						MSASFVPDHKASMVLFLDRVYGIENQDFLLH
1						VLDVGFLPDMRAAASLDTATFSTTEMALAV
1						NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
1						MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
						YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
1						TNHYERCWKYYCLPTGWANFGVTSEEELHL
				;		TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
		i i				AIAGALPPDYVDASYSSKAEKKATVDAEGNF
						DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
		-				KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
						EEKTEKKKTAKISOSAOTYDPREGYNPOPPDL
1 1						SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
						ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
						QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
						FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1 :	į					EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
			'			TPAKVLGSGGHASNKEKEMITSLFCKLAALV
		-				RHRVSLFGTDAPAVVNCLHILARSLDARTVM
					•	KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
						KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
]						QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
						TTKNTYVEKLRPALGECLARLAAAMPVAFLE
						PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
1 1						CPDIPVLERLMADIGGLAESGARYTEMPHVIE
]	•					ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
		ļ				CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
1		1				KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
j						AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
1						VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
1						AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
						NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
	.					RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
1		· /	[MCAPTDQDLITLAKTRYALKDTDEEVREFLH
	'	l				NNLHLQGKVEGSPSLRWQMALYRGVPGREE DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
						KKAVWHKLLSKORRRAVVACFRMTPLYNLP
1						THRACNMFLESYKAAWILTEDHSFEDRMIDD
						LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
] .		ſ				ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
						GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
						HTRGAAEMVLQMISACKGETGAMVSSTLKL
					-	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
]			J			IQALMQTCSVLDLNAFERQNKAEGLGMVNE
		i				DGTVINRQNGEKVMADDEFTQDLFRFLQLLC
		į				EGHNNDFQNYLRTQTGNTTTINIICTVDYLL
						RLOESISDFYWYYSGKDVIEEOGKRNFSKAM
						SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
						DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
		l				LDLQKDMVVMLLSLLEGNVVNGMIARQMV
		l				DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
		!				QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i.				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1]		residue of peptide	sequence	/=possible nucleotide deletion, \=possible
			[sequence		nucleotide insertion
<u> </u>		<u> </u>	-	sequence		QFLLSCSEADENEMINCEEFANRFQEPARDIG
						FNVAVLLTNLSEHVPHDPRLHNFLELAESILE
						YFRPYLGRIEIMGASRRIERIYFEISETNRAQW
						EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
Θ.						FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA
}						AEAGAEGAEGAAGLEGTAATAAAGATARV
						VAAAGRALRGLSYRSLRRRVRRLRRLTAREA
]						ATAVAALLWAAVTRAGAAGAGAAAGALGL LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
						SDEVHGEQPAGPGGDADGEGASEGAGDAAE
						GAGDEEEAVHEAGPGGADGAVAVTDGGPFR
						PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG
						VDGVEEELPPEPEPEPEPELEPEKADAENGEK
1						EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG
						ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN
ł						FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS GGSSGWGLGAGEEAEGDEDENMVYYFLEES
						TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
					-	LVIFKREKELARKLEFDGLYITEQPEDDDVKG
						QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG
						DIYGRERIAELLGMDLATLEITAHNERKPNPP
Ì						PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
						WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
						RTILSSYTHNGKQLVMTVGLLAVVVYLYTVV
1						AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV
						FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
					.	KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
						HNLANYMFFLMYLINKDETEHTGQESYVWK
						MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	Α	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
		. 1				LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD
502	1052	*	3000	1042	/47	YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL
						ASSDQSALASQSAGITGISTAPGPPFFFLNFEA
						GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR
				j		QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL
		- 1	ļ		1	VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP
]		J	J			HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPOIPGLKPSSCLRLLSSWDHRC
			ļ	j		APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
					ļ	L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR
			i			TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY
لــــا						QKEYYENQ*
505	1855	Α	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG
		1				NENTKLELRKVPPELNNISKLNEHFSRFGTLV
			ŀ		İ	NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL
	Į	Į				VQQPILPVVKQSVKERLGPVPSSTIEPAEAOS
		Į	- 1	İ		ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT
		Į	İ			LLVSTSAVDNNEAQKKKQEALKLQQDVRKR
	ļ	ŀ	- 1	į		KQEILEKHIETQKMLISKLEKNKTMKSEDKAE
]]	- [ĺ	i		IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI
	Í	[[ĺ	[KTKTQMQKELLDTELDLYKKMQAGEEVTEL
		ļ		ļ		RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA VHGRGRGRGRGRGVPGHAVVDHRPRALEIS
						VIOLOKOKOKOKO VPOHA V VDHKPKALEIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA
506	1856	A	3911	1952	919	VITFKTRAEAEAAAVHGARFKGQDLKLAWN KPVTNISAVETEEVEPDEEEQREIIIA DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLIRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLYYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEV\DLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKQRDKRNRHLGR
508	1858	Α	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	Α	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPKVTTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRMP\SP MAALILVADNAGGSHASKDANQVHSTTRRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGQARH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PPICHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIPPSRPDRSRNSNSLSR
513	1863	A .	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTIDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- ucnee	Met hod	SEQ ID NO; in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN VELGITISSPSSRKPGDPLVISDIKKGSVAHRT GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR YGGPLGVITISGTEEPVFDL*IISSLTKGGLAERT GAIHIGDRILLAINSSSLKGKPLSEAIHLLQMAG ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSAVIDTSYGTEGTVSFQASGYVNFNTYD WRSPKQRGSVLSPVTVKPRSQTYPDVGLSYED WDRSTASGFAGAAVDSAETEQEENFWSQALE DLETCGQSGILRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFFQQPSHGGNLETREPTVIL
514	1864	Α	3967	833	800	LEKQĞVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC
515	1865	Α	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	A	3977		1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV
517	1867	Α	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

nucleotide sequence In In In In In In In I	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Sequence	ոսշե-	peptide	ļ	in	лиcleotide	location	
uence	eotide	seq-	Ė	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
maino acid residue of peptide sequence propriet TeThreonine, Ve-Valine, We-Tryptophan, Ye-Tyrosine, Xe-Valine, We-Tryptophan, Ye-Tyrosine, Xe-Valine, We-Tryptophan, Ye-Tyrosine, Xe-Valine, We-Tyrosine nucleotide deletion, "possible nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide deletion and nucleotide de	seq-	uence	1	09/496		to last amino	
residue of peptide sequence y=Tyrosine, X=Unknown, **-Sico codon, *possible nuclocite delicion, *possible nuclotide insertion nuclotide insertion nuclotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide insertion Nucleotide Nucleoti	uence			914			
Peptide				ľ			
						sequence	
				ľ			
S20					sequence	<u> </u>	
1870			·				· ·
SIALLPRLYSNSWPQALLPRPKYLGLOT	520	1970	_	3000	002	608	
1871 A 4011 1346 1178	120	1070	^	3777	002	050	
PPTSASHYAGATGTHHHAWLSV	521	1871	Δ-	4011	1346	1178	
1872 A 4015 2 377				'''			
	522	1872	A	4015	2	377	
EYGPVSTWSALEGILAPPLEGVSACIGNCST				1022	_	1	
AL*ELTDOMTEDFLE*VLRE*VILY\$SDSMK							
L*NRREWDEAIKVLKEKQFLSKMVYPANLSF GNEGDITSPPAK	523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
SOMEGDITSPPAK							
1874							
1875							
RQQFTVLARMVLIS*PHDLPASASQSAGITGL SHCSWPTSSILS	524	1874	Α.	4020	1067	743	
SHCSWPTSSILS							
S25							
PPRLKO/FISHLSPPSINDYRRVPPCL VNFSIFF VETGSCQPC1QLIGSSNPASASQSAGIAGISH QGQPE*SFDIRFACVIAALRETFQCLCSASR VN NKIINRPTHPVESSF	526	1075		4001	701	261	
VETGSČQPCLQLLGSSNPPASASQSAGIAGISH QGQPE*SFDIRFACVIAALRETFQCLCSASRVN NKIINRPTHPVESSF	323	1875	A	4021	/81	331	
S26							
S26							
1876							
RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ	526	1876	Α	4024	80	341	
LHIHSSESQLHHSVKSPPSLSFRLM							
DVAVYFTTKEWAIMGVPAERALYRDVMLEN							- 1
1878	527	1877	Α	4026 ·	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
TGSSLSRNDWRAGWIGYLELRRYTYLS							DVAVYFTTKEWAIMG\PAERALYRDVMLEN
1878							· · · · · · · · · · · · · · · · · · ·
VEMGSAKŠVPVTPARPPPHNKHLARVADPRS							
PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSSIGSMRNRWKPN NSSKVLJGKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSEENVSELKEGAULGTGR\LLKTEGRA WEQQQD\HDKENQHFPLVES WEQQQD\HDKENQHFPLVES WEQGQD\HDKENQHFPLVES CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF HTEICT*MFIAVLFVVVKTWKQF RKYGSRNLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIGRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI S31	528	1878	Α	4028	1160	242	
DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSSIGSMRNRWKPN NSSKVLJGKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGA\ULGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES 529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF 530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/							
VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSS\GSMRNRWFP\							
QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGA\LLGTGRALKTEGRA WEQGQD\HDKENQHFPLVES 529							
ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES S29	ĺ						
NSSKVL\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\							
WEQGQNHDKENQHFPLVES	ł						
WEQGQNHDKENQHFPLVES							AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF 530							
ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF	529	1879	Α	4039	2	366	
HTEICT*MFIAVLFVVVKTWKQF	ļ			e-			
530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	ľ	l			İ)
RKYGSRINLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	520	1000		405-	250		
DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI	230	1880	A	4057	308	3	
VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	ļ	1					
531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/							
HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	531	1881	A	4061	50	278	
IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR		1001	"	1001	30	270	
T*KR T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	1						
532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/					j	A	
YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	532	1882	A	4069	19	368	
KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/			ĺ				
533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/	Ì	ļ			[
	533	1883	Α	4076	1	355	
KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC							KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC

cotide cence Sequence 9/44 corresponding feet f	SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
	eotide seq-	seq-		USSN 09/496	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
S34					residue of peptide		Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1884	 	<u> </u>	ļ		sequence		
DIQITTPKELEYTEKKENEL YESLAMIANIKÖR EMEMONIPTI.ITMKEELLDA TIMBERKOV EMEMONIPTI.ITMKEELLDA TIMBERKOV EMEMONIPTI.ITMKEELLDA TIMBERKOV VPENCEPVETREKCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA NKLISSYDYI.RESPVETILERCCIROJGELIISRI.NQAVA LWEQIKQIIQRITWVSPPAITLEWKRKVAQELISSYSTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQELISSYSTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQELISSYSTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQELISSYSTRM LEAGHSGRLEKTEDL.WILVSKOKYAVILLEKORYVSTVANIANIANIANIANIANIANIANIANIANIANIANIANIA	534	1884	A	4088	3	1931	IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL
VPENGEPYGTERIKCZIQQELISRINQAYA NKLISSYDYLIRESYGYTLERICLOSLEKSODVS							DLQITPKRLEYTRKKENELYESLMNIANRKQE
VHITSINYLKQILMAYHYEVFHISGSSYTRM							VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
LEAGHSGRLEKTEDL WLRVPKIGHARLARLS LESRSL QDVLLHRPKLGLGRGQYGVYL CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSIA VLLIMPELHRDLYTGLKAGLTLETRLQIALDV VEGIRFLHSQGLVHRDIKLKNVLLDKQRAKI TDLGFCKPEAMMSGSIVGTPHMAPELFTGK YDNSVDVVAFGILFWYICSGSVLKJEAFERCA SKDHLWNNYRRGARPELPYFDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKSI NSEQPNRGLDDST 1885 A 4090 2 417 ALIMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTRGLPSTLLAHIWALCDSKDCKLISKD HFALAFHLTIQKLIKGIDPLVLTTEKISPSNR ASLQKVTELTRPVCIIFKGTILWRITDSIWMK HNRKRIWLRA 536 1886 A 4102 569 829 DHQR*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKLPMTFFTEI*S*NVYRITKTQE*AKALLSKK EQNLESSHVLDFKYYRAK SQCNLESSHVLDFKYYRAK SQCNLESSHVLDFKYYRAK 537 1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRITHAQIKEHLWML HSI.GIDQQKTIE 1889 A 4111 268 IRPIPLIERSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS 539 1889 A 4111 268 I ASRPWGHSYP*FNQQEVDTLKRPLASSEI*MM I*KATIKKSPQPYRTAERTHKEDLVPPLW PLFPKIYREGTLPHSSYEASITL PEPGAGRAATPWGPLFWRGGSGGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHIGVRESGGAPQQPGRRRKRPR GRWREGCGAGGROVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPTTAVGEIDHYHLSEH IGALLIGEEVGAVFTPARVGRGGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHIGVRESGGAPQQPGRRRKRPR GRWREGCGAGGROVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPTTAVGEIDHYHLSEH IGALLIGEEVGAVFTPARVGRGGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHIGVRESGRAPQQPGRRRKPRR GRWREGCGAGGROVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPTTAVGEIDHYHLLSH MLKYYTTGRATTDEKEVLDFTSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSGGFLSLSK TALLNIVLRDSFAAPERDIFLALLNWCKHNSK ENHAEIMQAVRLPLIMSLTELLNVYRSGLLSP DAILDJIKVRSESSDMDLDWRYNDLDHG FSRHPIDDDCRSGEIKLGOPSINHVRILLHWDR DSRSYSYFIEVSMOELDWRYNDLDSYTCHQL BCGARSRNALLINGDTKNYDLDBGYTCHQL GGSGRVVCLAQPYMGELSGGFLAULDWCDANSY COMPTON TOWARD BCGARVCRUMCOCKHUMCHAUNTUNGENTHVAEI EGYSRSRNALLINGDTKNYDLDBGYTCHQL GGSGRVVCLAQPYMGELGSGTLULWDCDDRSY SAGNVVQLAQPYMGELGSGUTLUNDCDDRSY							VHITSNYLKQILNAAYHVEVTFHSGSSVTRM
CDNWGGHPCALKSVYPPDEKHWNDLALEF							LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS
							CDNWGGHFPCALKSVVPPDEKHWNDLALEF
TDLGFCKPEÄMMSGSIVGTPHIMAPELFTGK YDNSVDYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNYRGAPTERLPVTDECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS NSEQPNGLDDST NSEQPNGLDDST NSEQPNGLDDST STORT NSEQPNGLDDST STORT NSEQPNGLDDST STORT STORT NSEQPNGLDDST STORT							VLLIMERLHRDLYTGLKAGLTLETRLQIALDV
ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS NSEQPNRGLDDST	-	: 					TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK
1885							ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\
IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD	535	1885	A	4090	2 .	417	
HNRKRIWLRA Size						p•	IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV 537 1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML EVPVQRPVLYPQEQENEFSIGEFNEQVLRLM HSLGIDQQKTIE 538 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLITALPVS 539 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSFGPYRFTAFFSHTFKEDLVPILW PLFFKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSIPLRFFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MILKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALINWCKHNSK ENHAEIMQAVRLPLMSLTELLNVCRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEKLGQPSINHVRILLWDR DSRSYSYSTIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLINGDTKNYDWDSGYTCHQLG SGAVVQLAQPYMIGSIRVLLWDCDDRSY							HNRKRIWLRA
1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML EVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE	536	1886	A	4102	369	829	PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK
HSLGIDQQKTIE	537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML
VPRENITAWMNAIGLIITALPVS 539 1889 A 4111 268 I ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSSERDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							HSLGIDQQKTIE
I*KFAT\KKSPGYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRRPKKPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MILKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY	538	1888	Α	4109	141	314	
540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY	539	1889	Α	4111	268	1	I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
VRGHGGVRESGRAPQQPGRRRGRRPKKPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY	540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE
GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							
IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							GRWRREGCGAGGRGVCVAAWSQRSIAGNN
MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR
GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK . TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							
TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							GFPELEDSTSEYLCTILNIQNVCMTFDVASLY
DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							
MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY	1						
DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							MKYGAQVVKGELKSALLDGDTQNYDLDHG
WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							· ·
EGVSRSRNALLNGDTKNYDWDSGYTCHQLG SGAIVVQLAQPYMIGSIRVLLWDCDDRSY							WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF
							EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
	541	1891	Α	4146	282	778	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	6	1	D=Aspartic Acid. E=Glutamic Acid.
	I	noa	1	beginning	nucleotide	
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}	1	l	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}	ļ	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			l	peptide	•	/=possible nucleotide deletion, \=possible
1	1		1	sequence		nucleotide insertion
<u> </u>	-			sequence		
i	l	Í	l			HAESENFAFWQDMKWKNKFWGKSLEIVPVG
i	ŀ		l			TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
l .	1	ŀ				IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
	t					VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
		l				PTPGGQCIWKP
542	1892	Α	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA
• • •	.0,2		,	**	.55	QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR
•			i			, , ,
						INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT
	ļ					ANVSSENNTPRTSKTTFQLELSVKDAVYTVV
		<u>l_</u>				SSH
543	1893	Α	4153	678	11	TISYPOCLTOMYFLISFANVDTFLLPIMALDH
						YVAICSALQ*CSITTP/ELCOGLPVLA*AGSSLIS
	1		1	ľ		PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA
	!	I				CSHT*\NOHVFLGAVVLFLAPCALILVSYIRIA
1	!	1	1		1	, .
		1				AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV
		l	•			LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY
						SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP
						SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL
		1				LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS
1			1			LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG
		1	į.			
		1				LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE
						LKVGREGHVLPWQAHVVEF
545	1895	Α	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
1						DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
j		1				QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
		}			'	ELELSGEPAPGWQVLAGYTYTQARYLRDASE
1						ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	
340	1930	A	41/4	1232	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL
l l						QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
			1			FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA
						GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP
						GSHHASEGLAPGVPPAGGVSAQELTAPPOEG
						WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA
			1			RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS
i			ľ			, ,
į –			i			GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ
						PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL
						AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP
						HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS
						TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP
ļ l						ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/
						RCPPACSPAAASSFSFESQPCPSAPSKASPAPA
						AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL
						AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA
					İ	LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/
						PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
						AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ
						ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
						PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
						RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC
						LOPLHLRAAOPLDPCCSLSPPGPPLPVPALPS
						WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL
						PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ
						PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL
						TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
						MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP
		•				P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT
1						PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGPPAFQPPRFGCL*GPP GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG LHLPGGRTK
548	1898	Α	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVFHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	Α	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A	4192		1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQELKL LQEQDQELC'EILCMPHYDIDSASVPSLEELNQ FRQHVTTLRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATIL QKLLRQLEMQKSQNEAVCEG'LRTQIRELW DRLQIPEEEREAVATIMSGSKAKVRKALQ'LE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEEKQRAKLQKMLP KLEEELKARIEL WEQEHSKAFMVNGQKFME YVAEQWEMHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLL\ICITVCLSYI.PE AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSITFP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
552	1902	Α	4197	2	14302	ARPPPAPGSRQXQXAAPGAAAAAELRGAR
	17.02		,	-	11502	EPAPAKRRGTMADGGEGEDEIOFLRTDDEVV
						LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
						NSKNVPPDLSICTFVLEQSLSVRALQEMLANT
						VEKSEGQVDVEKWKFMMKTAQGGGHRTLL
						YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
i						VGLQEDTTGEACWWTIHPASKQRSEGEKVR
1						VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
1						QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
						GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
						VHARSLWRLETLRVAWSGSHIRWGQPFRLR
						HVTTGKYLSLMEDKNLLLMDKEKADVKSTA FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
						VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
						KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
•						TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
ļ						SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR
				ĺ		QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
l ,						VAGREAGESWKSILNSLYELLAALIRGNRKN
				İ		CAQFSGSLDWLISRLERLEASSGILEVLHCVL
						VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD
1			1			VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
					·	YELMVDHTEPFVTAEATHLRVGWASTEGYSP
						YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG
			1			CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
						RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
	•					RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
ĺĺ		- 1	1		•	KVEHSREYKQERTYTRDLLGPTVSLTQAAFT
		1	l	ł		PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
						KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
1	· j			1		ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
	1	ı	1	i		PSQEAMVDKLAENAHNVWARDRIROGWTY
1						GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS
		l				LREAVRTLLGYGYNLEAPDQDHAARAEVCS
	i	l	i	l		GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
						GDMRVGWSRPGCQPDQELGSDERAFAFDGF
	-					KAQRWHQGNEHYGRSWQAGDVVGCMVDM
	1	1	1	[ĺ	NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
1 1	ŀ	-	ł	į	ł	GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLOEGYEPFAVNTNRDITMWLSKRLPOFLOV
		İ	}			PSNHEHIEVTRIDGTIDSSPCLKVTOKSFGSON
	l	- 1	1			SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
1 1	·	- 1	1	İ		LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
]	1	1	İ		DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
	ļ	ĺ			1	RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
		- 1	- 1		1	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
		1		l	1	DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE
		- 1	-	l	1	SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
		ĺ		j	ĺ	VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
		1	- 1	ļ	ł	EHKNPVPQCPPRLHVQFLSHVLWSRMPNOFL
	1	- 1	- 1		l	KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
		ŀ		į	Į	RSVDILELTEQEELLKFHYHTLRLYSAVCALG
	1	ŀ	1	1	{	NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
	.	-			İ	AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
					ļ	EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF
	-	ı	1	ľ		SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE
						AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ŀ			residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				····		MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
						LEKELSVDDAKLQĞAGEEEAKGGKRPKEGLL QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
						VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
1						ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
ļ						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
						DSKKSSTLQQLISETMVRWAQESVIEDPELVR
						AMFVLLHRQYDGIGGLVRALPKTYTINGVSV EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
						LGDIMNNKVFYQHPNLMRALGMHETVMEV
						MVNVLGGGESKEITFPKMVANCCRFLCYFCR
						ISRQNQKAMFDHLSYLLENSSVGLASPAMRG STPLDVAAASVMDNNELALALREPDLEKVVR
						YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
						YLDFLRFAVFCNGESVEENANVVVRLLIRRPE CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
						GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
						SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS LIPLGDLVGVISIAFOMPTIAKDGNVVEPDMS
						AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
						EVGFLPDLRAAASLDTAALSATDMALALNRY
						LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
				j	j	MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
						YERCWKYYCLPGGWGNFGAASEEELHLSRK LFWGIFDALSQKKYEQELFKLALPCLSAVAG
						ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ
						PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE
			ŀ			KEIYRWPIKESLKTMLARTMRTERTREGDSM
						ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
						NVTLSRDLHAMAEMMAENYHNIWAKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
						QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
					,	YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
						SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
						RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKOGOF
			1			THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
			1			GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD
						KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
		i	l			SLEKLMEEIVELAESGIRYTQMPIIVMEVILPM
]			1			LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
						SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
						SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
		ŀ	ŀ			MVAEVFIYWSKSHNFKREEQNFVVQNEINN
.		1	1	İ		MSFLITDTKSKMSKAAVSDQERKKMKRKGD
						RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP
						AIRWQMALYKDLPNRTDDTSDPEKTVERVL
						DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK KAVWHKLLSKORKRAVVACFRMAPLYNLPR
	1	ĺ	[HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
			1	İ		KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT EKCKLEEDFLYMAYADIMAKSCHDEEDDDG
L						LACALLEEDT LIVIA I ADIMANSCHDEEDUDG

SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of peptide	hod	in	beginning nucleotide	nucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	(09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ļ				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ĺ	[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ	ļ	.)	peptide		/=possible nucleotide deletion, \=possible
	<u></u>	L		sequence	<u> </u>	nucleotide insertion
						EEEVKSFEEKEMEKQKLLYQQARLHDRGAA
	l	,		•	j	EMVLQTISASKGETGPMVAATLKLGIAILNGG
				ŀ	ŀ	NSTVQQKMLDYLKEKKDVGFFQSLAGLMQS
		ŀ				CSVLDLNAFERQNKAEGLGMVTEEGSGEKV
İ		ŀ		ļ	•	LQDDEFTCDLFRFLQLLCEGHNSDFQNYLRT
	·					QTGNNTTVNIIISTVDYLLRVQESISDFYWYY
				i	Ì	SGKDVIDEQGQRNFSKAIQVAKQVFNTLTEYI
						QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM
]		[QMKLSQDSSQIELLKELMDLQKDMVVMLLS
)		ļ	ļ	MLEGNVVNGTIGKQMVDMLVESSNNVEMIL
ļ						KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK RDFHKAMESHKHYTOSETEFLLSCAETDENE
						TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH
ļ			!			MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG
				1		SAKRIERVYFEISESSRTQWEKPQVKESKRQFI
ĺ					1	FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA
						OISESDLNERSANKEESEKERPEEOGPRMAFF
				}		SILTVRSALFALRYNILTLMRMLSLKSLKKOM
			1		ł	KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
}			,			ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL
}						ANMPDPTQDEVRGDGEEGERKPLEAALPSED
			•		}	LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP
						HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK
			1			EEEKEEKEETKSEPEKAEGEDGEKEEKAKED
						KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY
						QQKLLNYFARNFYNMRMLALFVAFAINFILL
i i]	1		:	FYKVSTSSVVEGKELPTRSSSENAKVISLDSS
						SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF
						FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI
						TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF
				•		VKRKVMDKYGEFYGRDRISELLGMDKAALD
	•	1				FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT
553	1903	A	4199	31	767	LPELNGRGAGLRRAEPSERGGGAERTOOVAA
333	1703	Λ.	4177] 31	707	LPLSHGHSHGGGGCRCAAER/VGAARGSAAC
						AYGLYLRIDKGRLQCLNESREGSGRGVFKPW
						ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK
[GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER
[EPDQ'IFSLNRDLTGELEYATKISRFSNVYHLSI
]						HISKNFGADTTKVFYIGLRGEWTELRRHEVTI
						CNYEASANPADHRVHQVTPQTHFIS
554	1904	Α	4200	1	961	GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL
j l				'		EICIKACKNLAYGEEKKKKCNPYVKTYLLPD
						RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA
				ļ		QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT
ļ ·				ļ		WDFEDSTTQSFRWHPLRAKADKYEDSVPQS
						NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL
! :						HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT
i i						LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV
						TPAQLRQSSLELTVWDQALFGMNDRLLGGT\
					1	RLGSKGDTAVGGDACSQSKLQWQKVLSSPN
555	1005	<u> </u>	4211	221	2410	LWTDMTLVLH
555	1905	Α	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH
						NHHENERR WQQERLHREEAYYQFINELNDE
				[DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE
						WLNTFRRTGNATRSGQNGNQTWRAVSRTNP
]	NNGEFRESLEIHVNHENRGFEIHGEDYTDIPLS
{					i	DSNRDHTANRQQRST\SPVARRTRSQTSVNFN
]						GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI
			L			

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	<u> </u>	1		residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence	i	nucleotide insertion
<u> </u>	 			sequence		GGAAGIPRANASRTNFSSHTNQSGGSELRQRE
						GORFGAAHVWENGARSNVTVRNTNORLEPI
						RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV
						QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
						TPFSNSRLVSRITVEEGEESSRSSTAVRRHPTIT
	Ì	1			ĺ	LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE
	i		ì			NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP
						LRRISENELVEPSSVALRSILRQIMTGFGELSSL
						MEADSESELQRNGQHLPDMHSELSNLGTDN
						NRSQHREGSSQDRQAQGDSTEMHGENETTQP HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH
) .		ł				FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN
1						SIDSELGKICSVCISDYVTGNKLROLPCMHEF
						HIHCIDRWLSENCTCPICRQPVLGSNIANNG
556	1906	A	4212	3	462	LOROROHPAAAPAVPVRCFTFCFTDIVIMPKR
						KSPENTEGKDGSKVTKQEPTRRSARLSAKPA
						PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK
						QEAGKEGTAPSENGETKAEEIHISRSTVNVST
						SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	Α	4213	774	507	ARRESCLTLQTSWGHRH\GPPRP\ANFVFLVET
						GFLHIGQAGHKLPTSGDPPASASQSARITGMS
558	1908	A	4225	3	1253	HRTWFLASFLIDSCKNFIVYKIMYTL TYRHAEREHPETSSATKVSYDYRHKRPKLLD
338	1908	A	4223	3	1255	GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE
						LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC
1						TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE
						SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK
		;				VDVKKTVDTFRVASSYSTERQMSHDLVAVG
						RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT
1						IIHQVKANYFPSPGITLHERFS\KMADIHKADV
1 1						NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE
						QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV
						EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ
						KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK
						KKVP
559	1909	A	4235	1	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL
1						LGLSHSPASASQVGGITGTQHHTGLIFGFLIET
						EFHHVGQAGLELLTSGDPPALAFQSAGITGVS
						HHAWLQVLNS
560	1910	A	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ
1						AALVNYSRLSEYAKIEGKKREMYELPVFCLA
						SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR
						LAELVIEVLQQNEEHHAEAFAWWSDLMVEH
						AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV
						R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN
						GSGTSEDLFWKLDALOTFIRDLHWPEEEFGK
1						HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK
				*]	TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP
]				•		KLICSMEMGQEFAKMWHQYHSKIDELIEETV
						KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS
			}			SFLSFTVKAASKYVDVPKPGMDVADAYVTF
				j		VRHSQDVLRDKVNEEMYIERLFDQWYNSSM
						NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY
						RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	Α	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI
"		•	''	1000	55 7	FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL
I			L			

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
иепсе		Ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i I				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		ļ		peptide	saquanto	/=possible nucleotide deletion, \=possible
İ		Ì		sequence		nucleotide insertion
				50,000		INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI
·					Ì	FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG
					1	DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ
1		l	l	ł		WGPRTNLETSKMKVLKFVAKVHNQDPKDW
			İ			PAQYCEALADEENRARPOPSGPAPSS
562	1912	A	4260	1	1498	MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF
302	1912	Α	4200	1	1470	WLHARLQKCFLSRGCGSYCAGAKASPLPGK
			1			MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ
		i	ł	}	}	WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR
1		ļ				i ·
		İ				SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV
			!	1		ASGETADVVQTAAEQSFAELGLGSYTPVGLI
			1	!		QNLLEFMHVDLGLPWWGAIAACTVFARCLIF
' '		1	1		[PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA
		ļ	1			GDHIEYYKASSEMALYQKKHGIKLYKPLILPV
		ŀ	1			TQAPIFISFFIALREMANLPVPSLQTGGLWWF
1		1		(QDLTVSDPIYILPLAVTATMWAVLELGAETG
						VQSSDLQWMRNVIRMMPLITLPITMHFPTAV
						FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR
						VVHDLDKLPPREGFLESFKKGWKNAEMTRQ
						LREREQRMRNQLELAARGPLRQTFTHNPLLQ
		<u> </u>				PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
563	1913	Α	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN
		Ì				THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK
1			·			HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L
J - 1						PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF
1						LIQENNNTNHTHSHTHTYTETLSFFLYICVNN
						DRMEWGKSVF
564	1914	Α	4270	3	368	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL
		1	[GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF
1						FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
						ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	Α	4288	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS
·						PPSALLAPTKPRALGTLRLYECSPELGTTMLP
					'	PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP
						GQTGASRTPRT
566	1916	Α	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
						GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
				i i		GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR
						VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ
						LLKKNGGIVMVTLSMGVLQCNLLANVSTVA
						DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E
]						DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR
						VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
						FHLGASEWTPRLLIWR
567	1917	Α	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
		1.7		-		DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
					i	WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
						MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
						VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
!						WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
						NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
						CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD
						INEAYVETLKHCFMMPQSLGVIGGKPNSAHY
						FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES
			·			FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
568	1010		4200	2012	1942	GAECCLGMTRKTFGFLRFFFSMLG
208	1918	A	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS
500	1010		4202	106		LMSVLIPKLPQLHGVRIFGINKY
569	1919	Α	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI
						CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT VTESKLEAEGKTKEKAREKERKKKS
570	1920	Α	4308	3		RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	À	4318		1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYVETL\KHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A .	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLK\ERM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSSIKRVLAITTVSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	Α	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Deadisted and	T A - ' '-' (A - A)- : C - C : -
NO: of	NO: of	hod	ID NO:		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	nou	l .	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	F - F		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	ценсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		I	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		· .	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
•	l	İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		İ	1	peptide		/=possible nucleotide deletion, \-possible
L				sequence		nucleotide insertion
577	1927	Α	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
1			i			ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
1						FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	Α	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
		l				SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
			ļ			SGWSRTPDLR
579	1929	Ā	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
1	1	1	1.000	_	1	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
]					CWPGWSSTPDLK
580	1930	A	4397	410	94	
300	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
j		l	!		}	\VFKKGI\IHILHELFQNKEEGAFPNS/FYEASFT
	Į.				i	LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
			<u> </u>		l	QLKSSDL
581	1931	Α	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
						RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
					1	RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
1	ĺ	ĺ	(RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
						LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
		l				VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
						DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
l l						SPE
582	1932	A	4424	194	449	VLYIRKKKRLEKLRHOLMPMYNFDPTEEODE
362	1932	^	4424	194	449	T COLL ELICED A GUOA ATQUO A AGGILTTI
		1		•	,	LEQELLEHGRDAASVQAATSVQAMQGKTTL
	1000					PS\QGPLQRPSRLVFT\DVANAIHV
583	1933 .	Α	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
						PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
						SAPPALLQDTSV
584	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
						APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
			1			APATQHSQAGPATGQAYGPHTYTEPAKPKK
			1			GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
			i l			ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
1 :						GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
			!			QGPHGKAAQGGAAGAAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
						SIFDDFAHYEKRQ
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLOTSDS
					- 00	FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
1						
]						INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
						FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
]	•					LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
[[[[PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
]					[ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
ļ ļ					ĺ	LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
į l				ľ	j	PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
]				İ		LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
1						SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
/						FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
						CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
1 1				ļ		NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
200	1550	^	1700	1,20	1470	
						CPANFCIII/DFLVETGFHHVGQASHELLTSGD
500	1020	- -	4407	000	-222	PPTSASQSAGITGMSYHTWFGES
589	1939	Α	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
j ļ				}	1	PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
; I				[1	REAAQAMYIDSYNSRGFMINGNRVLGPCALL
L1						PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence	_	nucleotide insertion VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ DTPNACATFNFLCHEGRVTGAALIPPPGGTSL TSLGQAAQ
590	1940	A	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS VPHQGGLPGPIRVAPSSAGQREASQGPPGR
591	1941	A	4495	1444	1116	IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT MEYYADTERNEIMSF\AGTWVELEAIILSKLM LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL EPWDSSCFPHPSSGV
592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL VCGGLSLLANAWGILSVGAKQKKWKPLEFL LCTLAATHMLNVAVPIATYSVVQLRRQRPDF EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR MWMVCWPVNYRLSNAKKQAGHTVMGIWM GSFILSALPAVGWHDTSERFYTHGCRFIVAEI GLGFGVCFLLLVGGSVAMGVICTAIALFQTL AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL GPFSLADTHLSDLPYTWGDRDSGGACVM
593	1943	Α	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC HFPASASQVAGTTHARHHTQLIF\AFLVENGL C
594	1944	A	4507		647	KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVR\DHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI
595	1945	A	4512	533	264	FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	A	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV
597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
598	1948	Α	4524	1	384	FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ	,	ļ		peptide		/=possible nucleotide deletion, \=possible
<u></u>				sequence		nucleotide insertion
						RRREMQSQSVMLALRRGDAVWLLSHDHDG
						YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG
						ASELL
599	1949	Α	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP
		•			ĺ	NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG
	1	!			[VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK
	i	i	}		}	HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI
	1000	ļ	1.50			PPPWLPKVLGLQA
600	1950	Α	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS
		1				DPPASASRVAGTTGARHHTQLIFVFLVETGFH
Ì		İ				\MLARDGLKLLTSSDPPASASQSSWDYRREPP
1	l	1]			RLANFFVFLVETGSRYVAQAGVQWLFTGAIP
		L				LLISTGVLTCSVSDLGRFTPP
601	1951	Α	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS
	İ	}				VGSPKAKEALNMLTWRAEQEGGMQFWVSSE
		1				SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE
		1				GIPIMRWLSRQRNSLGGFASTQDTTVALKALS
1		Į.	1			EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT
						HNRLLLQTAELADGTANGSV/SISANGFGFAI
		İ				CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV
						AVKENKDDLNHVDLNVCTSFSGPGRSGMAL .
		ŀ	ļ			MEVNLLSGFMVPSEAISLSETVKKVEYDHGK
						LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA
						SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD
	10.50					, VQRLPSL
602	1952	Α	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC
						VSGGLPKPANITFLSINMKNVLQWTPPEGLQG
						VKVTYTVQYFIYGQKKWLNKSECRNINRTYC
						DLSAETSDYEHQYYAKVKAIWGTKCSKWAE
					•	SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP
(EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT
						KSNRTWSQCVTNHTLVLTW\LEPNTLYCVHV
						ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK
					•	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK
						HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL
						NISADDSKISHQDMSLLGKSSDVSSLNDPQPSG
						NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\
1						EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI
						CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPOLODLDPLAQEHTDSEEGPEE
						EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE
						PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY
						LMQFMEEWGLYVQMEN
603	1953	Α	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL
""	.1703	12	7575	-	000	GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI
						LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL
						LHYAAETGNGEIVKYILDHGPSELLDMADSE
						TGETALHKAACQRNRAVCQLLVDAGASLRK\
1						TDSKGKTPOERAOOA\GDPDLAA/YTIESRON
						YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN
""	1754	^	4540	,	230	QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW
1						STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG
						QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS
				1		QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG
						FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA
						VKTVGSVVSSVALTGVLSGNGGTNVNMPVS
	İ					KPTSWAAIASKPAKPQPKMKTKSGPVMGGG
						LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP
						TITLE TENTINE TO TO THE TOTAL A LATE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
						QPPQ
605	1955	A	4553		2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE GPGLGALDRLRAHASAMGDEDLPGMAALQP HGVPGDGEGPHERGPPPASAPVGGTVTLRED SAKRLERRARRISACLSDYSLASDSGVFEPLT KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN LADYDSLSEMQLRWHSVQVFTSLNHQGRGR LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR TTAQLQAVERELAEERAKLEYTEEEVLEMER KEEQAEAISERSWQADSVDSGCSNCTQTSPPY PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL KVDKETNTEDLFLEEAASLVKERPSRRARGSP FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS STLPRKSPFVRNTLERTLRYKQSCRSSLAEL MARTSLDLELDLQASRTRQRQLNEELCALRE LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR EAERQTRQTKLDYRHEQAAEKMLKKASKEI YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL PADDV
606	1956	A	4555	3429		PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP VLLLQDSSGDYSLAHVREMACSIVDQKFPEC GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATTEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKHHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563	1	4499	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI QPGAFRRLRNLNTLLLNNNQIKRIPSGAFEDL ENLKYLYLYKNEIQSIDRQAFKGLASLEQLYL

NOT of NOT OF Petitide colded colded or corresponding in unclocation of the petitide colded or corresponding of the petitide of petitide colded in the petitide of petitide sequence petitide of petitide petitide of petitide petitide of petitide sequence petitide of petitide sequence petitide of petitide sequence petitide of petitide sequence petitide of petitide sequence petitide of petitide petitide of petitide petitide of petitide sequence petitide of petitide sequence petitide of petitide petitide of petitide petitide of petitide petitide petitide of petitide petitide of petitide petitide of petitide petitide petitide of petitide	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
location pt of irst amino acid residue of peptide sequence pepti		NO: of	hod		beginning	nucleotide	
uence unce control of the state of the state of the state of the state of peptide sequence	nucl-	peptide					
uence 914 and 16 fist animo acid residue of peptide sequence February F	eotide	seq-		USSN	location	, , ,	I=Isoleucine, K=Lysine, L=Leucine,
amino acid residue of peptide sequence peptide sequence of peptide	scq-	uence]	09/496	correspondi	to last amino	
residue of peptide sequence V=Tymsine, X=Uknkown, *=Siop codon, /=possible nucleotide delicito, \(\) \=p	uence]	l	914			
Popsible nucleotide deletion, \popsible sequence nucleotide insertion nucleotide insertion nucleotide insertion	1	1	ł				
Sequence nincleolide insertion HFNQETLIPPSSQHLPKLERIFLHNNRITHL VPGTPNHLESMKRLRLDSNTLHCDCEIL.WL. VPGTPNHLESMKRLRLDSNTLHCDCEIL.WL. DLLKTYAESGNAQAAALCEYPRRGGRSVAT TPEELINCERPRITSEPQDADVTSGNTVYFTCA AGOPPYEUWRNNNELSMKTDSKINLLDD GTLMIONTQETDQGTYQCMAKNWAGEWIT EVTLRYFGSPARPTYUQPONTEVLYCESVTL ECSATCHPPPRISWTRGDRTP.PVDPRVNTTYG GGLYIONVVQGDSGVACSATNNDSVHATZ GGLYIONVVQGDSGVACSATNNDSVHATZ FUNQALPQSTVTDQBVVAEGQTVDFQCEAK GNPPPVIAWTIKGGSQSVASCHATHAVISSGTLR SQVALHDQGQYECQAVNIIGSQKVAHLTW PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTTYFVGANVQLPCSSQGPE PPAITWNRDQVYTESGKPHISPEGFLTINDV PRVTPYFASPSDTLALAFTYPAGATGGSSQGARGGTVASSTVATSTVALAFTYPAGATGGSAVAGATGGSSQGARGGTVASSTVATSTVALAFTYPAGATGGSAVAGATGGAAVAGATGGAAVAGATGGAAVAGATGGAAVAGATGGAAVAGATGGAAVAGATGAAVAGATGAAVAGATGAATGA			ľ			sequence	
HFNOETLDPDSFOELPKLERIF-HNNRTITIAL VPGTFMHLESMKRLDSNTH-LOCELI-WIA DLLKTYAESGNAQAAAICEYPRRIGGRSVAT TPEELINCERPRITSEPODADVTSGNTVYFTCE AEGMPKPEUWLRNNNELSKKTDSRINLLDD GTLMIONTOETDOCIVOCMAKWA AGGVATT PEELINCERPRITSEPODADVTSGNTVYFTCE AEGMPKPEUWLRNNNELSKKTDSRINLLDD GTLMIONTOETDOCIVOCMAKWA AGGVATT EVTLRYFGSPARPTFVIQPONTEVL VCESVTIL ECSATCHPPPELSVTRGORTPT-PVDPEVNTTE GGLYQNVVQGDSGEYACSATNNIDSVHAT/FIVQALPQFTVTPQDRVVPGGTVDFQCEAK GNPPPVAWTKGGGQLSVDRTHLVISSGTLR SGVALHDQQYECQAVNIIGSQKVAHLTVV PRYTFYHASTPSDTTVEVGANVQLPCSSGGE EPAITWNICDQVPTCSSGFHISFEGFLTINNDV GADADGAYECVARTISGSASVSMVLSVNPV VSRNGDPFVATSIVESAIATVDRAINSTRTHLF DSRPRSPDLLALFRYPRDFYTYEQRA GEEI ERTLQLIQEHVQHGLMVDLNGTSYTHYNDLV PQVLNLANLSGGTAHREVNNCSDMCFHQK RIHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRILYNGHALPMRLVSTTL GTETVTPDEQTHMLMQWGGFLDHDLDSTV VALSQARSDQQHGSNVCSNDPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMNN VYRREGNVCSNDPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMNN VYRREGNVCSNDPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMNN VYRREGNQULSTURLANDSTRTHLFF DSRARSGARCMFFVRSSPVCGSGMTSLLMNN VYRREGNQULSNVGSMPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMNN VYRREGNQULSNVGSMPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYRREGNQULSNVGSMPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYRREGNQULSNVGSMPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYRREGNQULSNVGSMPFCSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYRREGNQULSTAFARSIND LASHGARATAAAFRGGTTLVNPLLLFGLL ENGPAIN VALDLAANINGARGMFTGFFTGCMSTAGGFTGCFTGCFTGCFTGCFTGCFTGCFTGCFTGCFTGCFT							
VPGTYNHLESMKRIR,DSNT1.HCDCBILWLA DLLKTYAESGNAQAAAICEYPRRGQRSVAT TPEELNCERPRISEPQDADVTSGNTYYFTCA AEGMPYEUW.RNNNELSMKTDSKINLLDD GTLMIONTGETDQGIYQCMAKNWAGEWIT EVTR.YFGSPARPTYUPQDNTEVLYCESVTL ECSATCHPPPRISWTRGDRTP.PVDPEVNTTF GGLYQNVVQGDSGEVACSATNNIDSVHAT/ FIVQALPQFTVTPQDRVVBCQTVDFQCEAK GNPPVLAWTKGGSQLSVDRRHLVLSSGTLR SGVALHDQQYECQAVNIIGSQKVAHLTVV PRVTPYASIPSDTTVEVGANVQLPCSSQGE EPAITWNLDGVQVTPESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPP VSRNGDPPVATSIVFESGKFHISPEGFLTINDV FOYLNLANLSGCTAHRKVNNCSDMCFHGK RTHDGTCNNLQHWMQASLTAFERLLSSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLL GTETVTPDEQFTHIRMMOWGGFLDHDLDSTV VALSQARSSDQOHCSNVCSNDPCFSVNIFPP DSRARSGARCMFFYSSEVCGSGMTSLLMN VYPREGINQLTSVIDASNVYGSTEHEARSILU LASHRGLRQGIVGSGKPLIPFAGTPTECK RDENSEPPFFLAGDHRANEQLGLTSMFTLW FRHNRIATELLKAUPHWDGDTIYPCTRKVV AEIQHITYQBWLPKLLGEVGMRTLGEYHGTV PGINAGIFPAFATAARAFRGHTLNDPHLLPGAL ENFQPIAQDHLPHKAFFSFFRIVNEGGIDPLL RGLFGVAKMVPSQLLNTELTERLESNAFH VALDLAANINGKRAPRGHTLNDPHLLPGAL ENFQPIAQDHLPHKAFFSFFRIVNEGGIDPLL RGLFGVAKMVPSQLLNTELTERLESNAFH VALDLAANINGKRAPRGHTLNDPHLLPGAL ENFQPIAQDHLPHKAFFSFFRIVNEGGIDPLL RGLFGVAKMVPSQLLNTELTERLESNAFT VALDLAANINGKRAPRGHTLNDPHLLPGAL ENFQPIAQDHLPHKAFFSFFRIVNEGGIDPLL RGLFGVAKMVPSQLLNTELTERLESNAFT VALDLAANINGKRAPRGHTLNDFTRHV DGRUMPKFRIGHTAGRSCALLGCGEPRIX LFSTRDASGTNDFQRVCSWEMQKTITDLR TYPAQUECCCEPTTATATDSHFFLAYAGRSTLANTSW KDACTICECKDGQVTCFVEACPPATCAPPYN LRCHLESRUSTTECVDAGGESHANNTSW KDACTICECKDGGVTCFVEACPPATCAPPYN LRCHLESRUSTTECVDAGGESHANNTSW KDACTICECKDGGVTCFVEACPPATCAPPYN LRCHLESRUSTTECVDAGGESHANNTSW KDACTICECKDGGVTCFVEACPPATCAPPYN PGACCPVCLQKRAEEPP GECCTVLQKRAEEPP GECCTSVRSNSGLSIGTVPQSSCSPGGGG LGVDAW RAPRIKRNVYTSFSAGSAGGAGGTGVSH RAPPRIKLNNVYSFSAAVTVCLNYSLAMSSTIL KSHWLSSCS EKGVSSAMPCTLACH RESEDULTSSCSSASAF	ļ		<u> </u>		sequence		
DILLETYAESGNAQAAAICEYPRRIQGRSVAT TPEELICERRITISEDQADAYTSGNTYFTCE AEGHPKPEUWLRINNELSAKITDSRIALLDI GTLMIGNTGETDQCIYQCAMAXWAGEVATG EVTLRYFGSPARPTPVQPQNTEVLVGESVTL ECSATGHPPRISWTRGDRTPLPVDPRVNTIPY GGLYQNVVQGDSGEVACSATNINISVBATA FIIVQALPQFTVTTQDRVVIEGQTVDFQCEAK GNPPPVLWTKGGGGQLSVDRRHLVLSSGTLR SGVALHDQQYECQAVNIIGSQKVVAHLTVV PRVTPYFASIPSDTTLEVGANVQLPCSSQGEP EPAITWNRDQVYTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSWAWLSVNVP VSRNGDPFVATSIVEALATVDRAINSTRTHLF DSRPRSPDDLLALFRYPNDPYTVEQARAGEIF ERTLQLIQEHVQHGLMVDLNGTSYHYNDLV PQVLNLIANLSGCTARRVNNCSDMCFHOK RIHDGTCNNLQHPMGASLTAFERLLKSVV ROGRIPPRGINPPRILYNGHALPMRLVSTTL GTETYTPDEQFTHMLMQWGGPLDHDLDSTV VALSQARFSDQHCSNVCSDPFCFSVMIPP DSRARSGARGMFVRSSVCGSGMTSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIND LASHRGLRQGIVQSKEPLLFFATGPTECK RDENESPPCTLAGDHRANGQLGLTSMHTLW FREHNRATELLKLNFHWDGDTTYPCTRIVK AEQHITYQHWLPKLLGEVGMRTLGEYHGVY PGINAGIFNAFTAAFRGHTLVNPYLLPGLL ENRQPJAQDHLPLHKAFFSFRINVBGGDPLL ENRQPJAQDHLPLHKAFFSFRINVBGGDPLL ENRQPJAQDHLPLHKAFFSFRINVBGGDPLL ENRQPJAQDHLPLHKAFFSFRINVBGGDPLL ENRQPJAQDHLPLHKAFFSFRINVBGGDPLL LRVWQDCCEDCTREPRISEKLRILYGSTLAND LPPALVVEDLVPGSRLGPTLMCLLSTQFKRL DGDRLWYENPQVFSPAQLTINGLTGLFSMHT VALDLAANIQRGRDHGIPPYHDYRVYCNLS AHFTEDLKSBIKNPERIEKLRILYGSTLAND LPPALVVEDLVPGSRLGPTLMCLLSTQFKRL LRVWQDCCEDCTREPRISEKLRILYGSTLAND LPPALVVEDLVPGSRLGPTLMCLLSTQFKRL LRVWQDCCEDCTRETRGQPHASYHFRGRSLL DGDRLWYENPQVFSPAQLTQKGTSLARIC NADMTRVGSDVFRVAEFPHGYGSCDEIPRVL LRVWQDCCEDCTRETRGQPHASYHFRGRSLL FSYQEDKPTKKTRPKIPSVGRQGEHLSNSTS AFFSTRDASGTNDPGRVCSWEMQKTTIDLR TQKKLESRLSTTECVDAGGESHANNTKW KDACTICECTDAGGESHANNTKW KDACTICECTDAGGESHANNTKW KDACTICECTDAGGESHANNTKW KDACTICECTDAGGEGGGGGVRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKKRLGINLLGGEIGGKQVIRVSCCG LGYGNYKRGLGINLGGRAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGA			İ]			
TPEELNCERPRITSEQDADVTSCNITVYTER AEGPKPKPEIWLRNNELSMKTDSELNILDD GTLMIQNTQETDQGIYCCMAKNWAGEWLF EVTLRYFGSPARPFTVIQQNTPELVCESYTT EVTLRYFGSPARPFTVIQQNTPELVCESYTT EVTLRYFGSPARPFTVIQQNTPELVCESYTT GGLYQNVVQGSGEXFACSATNIDSVHAT, FIIVQALPQFTVTPQRVVIEGOTVDPQCEAK GNPPPVLAWTKGGSQLSVDRRHLVLSSGTTL SGVALHDQGYQFQQAVMIGSGKVAHLTV PRVTPVFASIFSDTTVEVQANVQLPCSSGE EPAITVNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVMIP VSRNGDFPVATSIVEALATVDRANSTETHLF DSRRSPNDLLALFRYPRPPYTVEQARAGEIE ERTLQLIQEHVQHGLMVDLAISTAYNNDLV PQYLNLLANLSGCTAHRRVNNCSDMCFHQK RIHDGTCNNLQHPMVGASLTAYNNDLV PQYLNLLANLSGCTAHRRVNNCSDMCFHQK RIHDGTCNNLQHPMVGASLTAYNNDLV PQYLNLLANLSGCTAHRRVNNCSDMCFHQK RIHDGTCNNLQHPMVGASLTAYNNDLV PQYRANLANLSGCTAHRRVNNCSDMCFHQK RIHDGTCNNLQHPMVGASLTAYNNDLV PQYRANLALSRYSHVOLD ERFRIPTION OF THE CONTROL OF THE CONTROL ERFREYNOLL STRONG THE CONTROL ERFRYNTEGRPHELLYSYY ENGRITHERGRPHILYSYGGABGEIE LASHRGLARQGVQRSKCHLPTGPTECN RDEMESPIPCTLAGDHRANEQLGLTSMHTLW PREINRATELLKLNHPWDGDTHYTPGRPTCA AEQHITYOHWLPKILGEVGMRTLGFYHGYY PGINAGIPNAFATAAFFFCHIVPEGGIPLL ROGHTYOHWLPKILGEVGMRTLGFYHGYY PGINAGIPNAFATAAFFFCHIVPEGGIPLL ROGHTYOHWLPKILGEVGMRTLGFYHGYY PGINAGIPNAFATAAFFFCHIVPEGGIPLL ROGHTYOHWLPKILGEVGMRTLGFYHGYY PGINAGIPNAFATAAFFFCHIVPEGGIPLL ROGHTYOHWLPKILGFYHORYSTAALFFYHTVRLLSTOLL ENGPOPAGOMPLPHIKAFFFFRIVPEGGIPLL ROGHTYOHWLPKILGFYHORYSTAALFYHTROR AHTTELLANLSTOLL ENGPOPAGAMTAGCTPCTRAGTSTAATAFFFTHYTPLGFTCH AEGHTYOHWLPKILGFYHTTSTTAACUST AAHTEBLANERIKRYPEIRSKLKRLYGSTIND LFALVYEDLVFSRAQLTQRTSTYPHFRGRRSL LFAVYEDLVFSRAQLTQRTSTYPHFRGRRSL FSYCGGKGCLDSEEDYYTTPCKYDVFX AFSTRSDASGATDDFQRVCSWEMQKTITUL TQKKKLESRLSTTECVDAGGESHANNTKUK KDACTICECKDGQVTCFVACPPATCAVPVN PGACCPVCLQKRAEKP TQKKKLESRLSTTECVDAGGESHANNTKUK KDACTICECKDGQVTCFVACPPATCAVPVN PGACCPVCLQKRAEKP FSFLCGGVGRCLDSEEDYYTTPCKYDVFX FKIRTAATDDNIHEATHADERGRILHANTICR GDDAW UNVOCGCGCTFTLTGGSKVAACGLOSEEDYTTTAACUST GCDDAW UNVOCGCGCTTAADDFRANTAMIPTER HGSDICTSWPRPIGGSLHANTKUK ROGHTHAATDDNIHEATHADERGRILHANDHAATHADERGRICHANTSTAA FSTRSDASGATGTONTSTAAT FFFTETESTRAQAQGVQWRDLGSLQAPPFGFF FSCCLSLPSSWDVRRPLRANFFVILVETGF HRYSDGGLLLTYSSC							l
ABGNPKPEIWLRNNNELSMKTDSRLNLLDD GTLM(MONTOSTDOGIYOCMAKGEWITG EVTLRYFGSPARPIFVIQQENTEVLVCESY) EVTLRYFGSPARPIFVIQQENTEVLVCESY EVTLRYFGSPARPIFVIQQENTEVLVCESY EVTLRYFGSPARPIFVIQQENTEVLVCESY GGLYIQNVVQGDSGEYACSATNNIDSVHAT/ FIVQALPGFTVTPQDRVVLEQQTVDPQCEAV GNPPVLAWTKGSQSLSVDRRHLVLSSGTLR SGVALHDQGYECQAVNIGSQKVVAHLVT PRVTPYFASIPSTITVEVGARNQLPCSSQGEP EPAITWNKDGVQVTESGFHISPEGFLTINDT OPADAGRYECVARNTIGSASYSMVLSVNVP VSRNGDPFVATSIVEALATVDRAINSTRTHLY DSRRSPNDLLALFRYPRDPYTYFQARAGEI ERTLQLIGEHVQHGLMVDLNGTSYHYNDI VSRNGDPFVATSIVEALATVDRAINSTRTHLY DSRRSPNDLALFRYPRDPYTYFQARAGEI ERTLQLIGEHVQHGLMVDLNGTSYHYNDI VSRNGDPFVATSIVEALATVDRAINSTSHANDLY VSRNGDPFVATSIVEALATVDRAINSTSHANDLY VSRNGDPFVATSIVEALATVDRAINSTSHANDLY VSRNGDPFVATSIVEALATVDRAINSTSHANDLY VSRNGDPFVATSIVEALATVDRAINSTSHANDLY VSRNGDPFVATSIVEALATVDRAINSTSHANDLY ERTLQLIGEHVQHGLMVDLNGTSYHYNDI VSRNGDPFVATSIVEALATVDRAINSTSHANDLY ERTLQLIGEHVQHGLMVDLNGTSYHYNDI VSRNGDPFVATSIVEALATVDRAINSTSHANDLY ERTLQLIGEHVGHGLMSVHAULT ERTLYTDEOFTHMLMQWGQFLDHDLDSTY VALSQARESGAD(CSNVCLSUPETVMIPP DSRRARSGARCMFFVRSSPVCGSGMTSLLMN VYPREQINGLTSYTDASTSHANDLY ERTLYTTE ERTLYTTL ERTLYTTE ERTLYTTL ERTLYTTL FREHINRIATELLKLINHWDGDITYYETRILV AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLGEYHGVT AEQHTYQHW_PRILGEVGMRTLLGEYHGT AEQHTYQHW_PRILGEVGMRTLGEYHGT AEQHTYQHW_PRILGEVGMRTLLGTATTSHANDLT AEQHTYQHW_PRILGEVGMRTLLGTATTSHANDLT AEQHTYQHW_PRILGEVGMRTLLGTATTSHANDLT AEQHTYQHW_PRILGEVGMRTLLGTATTSHANDLT AEQHTYQHADATAATAATAATACHTAATAATAATAATAATAATAATAATAATAATAATAATAA	1		l			ļ	
GILMIONTOETDOGITYCCMAKNIVAGEWRT, EVTLRYRGSRAPPTFVIOQOTTEVCSEVTIL ESCATGHPPRISWTRGDRTPLPVDPRVNITP GGLYYONVVQGDSGYACSATNIDSVHATP GGLYYONVVQGDSGYACSATNIDSVHATP GGLYYONVVQGDSGYACSATNIDSVHATP FITVQALPQFTVTPQDRVVIEQQTVDFYQCEAK GNPPPVLAWTKGGSQLSVDRRHLVLSSGTIL SGVALHDQGQYEQQAVNIIGSGKVVAHLTV PRVTPVFASIPSDTTVEVQANVQLPCSSQGE EPATTWNKDGVQVTESGKFIPSPGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPI VSRNGDPFVATSIVEALATVDRATRTHLF DSRRRSPNDLLALRTYPRDPYTVEQARAGEI ERTLQLIQEHVQHGLMVDLNGSTRYNDLV PQYLNLANLSGCTAHRRVNNCSDMCFHQK- RIHDGTCNNLQHPMWGASLTAFERLLKSVY ENGMTPRGINPHRLINGGALAHPMRLVSTILL GTETVTPDEQFTHMLMQWGQFLDHDLDST VALSQARSSDGOHCSNVCSNDFCFSVMIPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYPREQNNQLTSYDASNVYGSTHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPFTEC RDENESPIPCTLAGDHRANEQLGITSMHTLW FREINRIATELLKLNPHWDGDTTYFETRIVC AEQHTYQHWLPKLIGGEWGRIGGTSMHTLW FREINRIATELLKLNPHWDGDTTYFETRIVC AEQHTYQHWLPKLIGGEWGRIGGHSNINGLBYGHGY PGINAGIFNAFATAAFRGHTLVNPLLJEDL ENGPPAQOHLPLHKAFPSPFHVGGGDPLL ENGPPAQOHLPLHKAFPSPFHVGGGDPLL ENGPPAQOHLPLHKAFPSPFHVGGGDPLL ENGPPAQOHLPLHKAFPSPFHVGGGDPLL DGDRLWYENROVFSPAQLTGHGYSLARILCI NADNITRVQSDVFRVABFPHGVGSCDEHSYN LRVWQDCCEDCRTRGGPNAFYHFRGRRSL FSYQEDKPTKKTRPRKIPSVGRQGEHLSNTDLR AUFTSDAGSGTINGCTSLARILCI NADNITRVQSDVFRVABFPHGVGSCDEHSYN LRVWQDCCEDCRTRGGPNAFSYHFRGRRSL FSYQEDKPTKKTRPRKIPSVGRQGEHLSNTDLR AUFTSDAGSGTNTDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTTECKDGVYTGVAGGSHANNTKWK KDACTTECKDGVYTGVAGGSHANNTKWK KDACTTECKDGVYTGVAGGSHANNTKWK KDACTTECKDGVYTGVAGGSHANNTKWK KDACTTECKDGVTGVAGGSHANNTKWK KDACTTECKDGVTGVAGGSHANNTKWK KDACTTECKDGVTGVAGGAGHLSWTDLR GLNQCMSGINHEAYHEVPYTTSTTLAKQLST YKIRTAFOKHTAADERGRILTGCNKCGG LGVONYKKRIGNILGGPLGGKQVIRVCGGNEFNK GLNQCMSGINHEAYHEVPYTTSTTLAKQLST YKIRTAFOKHTAADERGRILTGCNKCGG LGVONYKRRIGNILGGPLGGKQVIRVCGRWHILL VEKVLNSKTIRSNSGLSIGTVAGSSCGGGGG GGPDAW FFFETEESSNVAQAGVQWRDLGSLQAPPGFF FFSCLSLPSSWDYRPPLRPANFFVFLVETGF HRSBDGLDLLTSNGDPFASASQSAGITGVIRVS RAFRINLRNIVTSAVTYCATVISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRCVISAHCCTLCLPSSSDSASAAPRVARTT							
EVTLRYEGSPARPIFVIOPONTEVLYGESYL ECSATGHPPPRISWTRGDRIPLPDPRVNITP GGLYIQNIVQGDSGEYAGSATNNIDSHAFT FILVQALPGETYPTQDRVVIEGGTVPPCQEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLR SGVALHDQGYEQAVNIGSQKVVAHLI- FILVQALPGETYPTQBRVVIEGGTVPPCQEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLR SGVALHDQGYEQAVNIGSQKVVAHLI- FILVQALPGETYPTQBRVVIEGGTVPPCQEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLR SGVALHDQGYEQAVNIGSQKVVAHLIV PRYTPYFASISTIVEAGATVDRANISTSHVIVPI GPATTWNKDGVQTESGFHISPEGFLTIND GPADAGRYECVARNIGSASVSWLSVNIVPI VSRNGDPFVATSIVEAGATVDRANISTSHVIVPI DSRRSPNDLLAFRYPRDYTYPCQARAGEII ERTT_QLIQEHYQHGLMYDLNGTSYHYNDI PRYTLALIFRYPRDYTYPCQARAGEII ERTT_QLIQEHYQHGLMYDLNGTSYHYNDI FILVGENTPRGIPPHLYYGHALPRILVSTTL GTETYTPDEQFTHMLMQWQQFLDHDLDST FILVGARAFSGOGCSMYSSDPCFSVMIPP DSRARSGARGMFYRSSPVGSGGMTSLLMN VYPEGDNQLTSYDANNYQSTEARSIND VYPEGDNQLTSYDANNYQ		ŀ	1				L
ECSATGHPPRISWTRGDRTPLPVDPRVNITP GGL YIQNNVQGDSGEYACSATNNIDSVHATY GGL YIQNNVQGDSGEYACSATNNIDSVHATY FIVQALPOFTVTPQDR VVIEGQTVDPGCEAK GNPPVLAWTKGGSQLSVDRRHLVLSSGTLR SCVALHDGGQYECQAVNIGSQKVVAHLTVV PRVTPVFASISPDITTVEVGANVQLPCSSGGEP EPAHTWNKDGVQVTESGKHISPBGFLTINDV GPADAGRYECVARNITGSASVSMVLSVNVPI VSRNGDPFVATSIVEALATVDRAINSTRTHLP DSRPRSPNDLLALFRYPRDPYTVEVGANGEIF ERTLQLIQEHVORIGLMVDLANTSTHTHLP DSRPRSPNDLLALFRYPRDPYTVEVGARGEIF ERTLQLIQEHVORIGLMVDLANTSTHYNDLV PQYLALIANLSGCTAHRRVINCSDMCFHQK RIHDGTCNALOHPMWGASTAFERLLKSVY EGGFITPRGINPHRLYNGHALPMPLVSTTL GTETVTPDEQFTHMLMQWGGFLDHDLDSTV VALSQARRSDQCHCSNVCSNDPPCFSVMIFP DSRASSGARCMFFVRSSPVCGSGMTSLLMN VYPREQINQLTSYIDASNVYCSSTEHEARSIKD LASHRGLLRQGIVQRSGKPLLPFATCPPTECK REDENSPPCTLAGDHRANEQCLSTSMHTLW FREHNRIATELLKLNPHWDGDITYYFTRIV ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWDGDITYYFTRIV ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWDGDTYYFTRIV ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWPDGTYPYCYCNLS ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWPDGTYPYCYCNLS ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWPDGTYPYCYCNLS ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWPDGTYPYCYCNLS ARIGHTYQHULPKLIGEVGMRTLGEYHGY FREHNRIATELLKLNPHWPDGTYPYCYCNLS AAHTFEDLKNEKNPEIREKLKRLYGSTLNID LEVALVVEDLVPGSRLGFTHMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI DGDRLWYENPGYFSAQLITMCLISTGFRLI TQKKLESRUSTTECVDAGGESHANNYLWK KDACTICECKEGQVTCTPVAGFVAGGENENNI LEVALVOCCCGGFTPLTTQSGVLACGLNEFNKI GLOCK GGGFDAW TRYNOCCCCGGFTRLTTGGRVLACGC LETTAL TOTAL ARIGHT VEKANDAL TOTAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT VEKANDAL ARIGHT FSFLCGVSGRUGHGSLAGGRUAMTPTERF	ł		l				
GGLYIQNIVQGDSGEYACSATNINIDSWIATA FIIVQALPOFTVPTORDR VYBEGOTVDPOCEAK GNPPVIAWTKGGSQLSVDRRHLVLSSGTLR SGVALHDGGQYECQAVNIIGSQKVVAHLTVI PRVTPVPASIPSDTTVEVGANVQLPCSSQGEP LEPAITWNKDGVQVTESGKFHISPEGETLINDV GPADAGRYECVARNITGSASVSMVLSVNVPI VSRNGDPFVATSIVESAIATVDRAINSTRHLF DSRPRSPNDLLALFRYPRDPTVTVEQARAGEII ERTLOLIQEHVQHGLMADLANGTSYHYNDLV PQYLNLLANLSGCTAHRRYNGSMCFHAVE RINDGTCNNLQHPMWGASLTAFERLLKSVY ENGRYTRGRIPHRLYNGHALPMRUSTILL GTETVTPDEGFTHMLMQWGQFLDHDLDSTY VALSQARFSDOQHCSNVCSNDPPCFSWIPPI DSRARSGARGHFVRSSPVCGSGMTSLLMN VYPRRGINQLTSVIDASNVYGSTEHEARSIRD LASHIGLLRGGIVQRSGEHLPFATGPFTECK REDENSSPCCLAGDHRANEQLGLTSMHTLW FREINRIATELLKLNPHWDGVGSCHERNYCS REHERLIKLLNPHWDGVGVGSGMTSLLMN VYPRRQINQLTSVIDASNVYGSTEHEARSIRD LASHIGLLRGGIVQRSGEHLPFATGPFTECK REDENSPCCLAGDHRANEQLGLTSMHTLW FREINRIATELLKLNPHWDGVGSGMTSLLMN VYPRRQINQLTSVIDASNVYGSTEHEARSIRD LASHIGLLRGGIVQRSGEHLPFATGPFTECK REDENSPCCLAGDHRANEQLGLTSMHTLW FREINRIATELLKLNPHWDGVGSTEHERVING AEIGHTYQHWLPKILGEVGMRTLGEYHGVT PGINAGIRNAATAARROCHTLVNPLLIFGLI ENFQPIAQDHLPLHAFFSFRIVNEGGIDPL RGLFGVAGKMRVPSQLLINTELTERLFSMHT VALDLAANIQRGDHGIPPTHDYFTSWTXAN AAHTFEDLKNEIKNPERREKLKELYGSTIND LFPALVVEDLVPGSRLGFTLMCLLSTOFKRIL DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITEVQBOVFRVAEPFRIVNEGGIDPLI REVENDOCEDCETTGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNITS AFSTRSDASGTNDPGRVCSCDEHRVXI LFVWQDCCEDCETTGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNITS AFSTRSDASGTNDPGRVCSCDEHRVXI LFVWQDCCEDCETTGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNITS AFSTRSDASGTNDPGRVCSCDEHRVXI GLNQCMSGINMEAYHEVPTSTTLAKQLSS YKRTIAPGKTHTAAIDERGRLLTFGCNKCGG LGVGMYKKRLGINLLGGPLGGKQVIRVSGD EFTIAATDDINFFAWGNGGGALAMIPTERFR HGSDICTSWPRPIGSSLHHVPDLSCGWHTIL VEKVLNSKTIRSNSSGLISGLAMPTYGLVFTGF HGSDICTSWPRPIGSSLHHVPDLSCGWHTIL VEKVLNSKTIRSNSSGLISGLAMPTYGLVFTGF HGSDICTSWPRPIGSSLHHVPDLSCGWHTIL VEKVLNSKTIRSNSSGLISGLAMPTYLVFTGF HGSDICTSWPRPIGSSLHHVPDLSCGRUFTNL GGPDAW RARRINLRNVTYSFAVTYCLNYTSLAMSSTL KLSFHVLSGS 609 1959 A 4567 1 412 PFPFETESRSVAGGOVQWBLGSLQAPPGF FPSCLSLPSSVDYRRPPLRPANFFVTLVETGF HRSRDGLDLLTXSGDPPAGSAGRATHVARTT KLSFHVLSGS 608 1960 A 4570 697	İ						1
FIIVQALPOFTVTPODRVVIEGGTVDFOCEAK GNPPPVIAWTGGGGUS.VPAILUSGGTLAR SGVALHDQGQYECQAVNIIGSQKVVAHLTV PRVTPVFASIPSDTTVEVGANVQLPESSGGEP EPAHTWNKDGVQVTESGKEHISPBGFLTINDV GPADAGRYECVARNITGSASVSMVLSVNVPI VSRNGDPFVATSIVEALATVARINSTRTHLEF DSRPSSPNDLLALFRYPRDPTVFVGARGEIF ERTLQLIQEHVORIGIMVDLAGTSYHVNDLV PQYLNLIANLSGCTAHRRVNNCSDMCFHQK RITHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINFHRLYNGHAPPREVSTTLY GTETVTPDEQFTHMLMGWGQFLDHIDLDSTV VALSQARFSDOQHCSNVCSSDPFCFSVMIPPI DSRARSGARCMFFVRSSPVCGSGMTSLIMN VYPREQNQLTSYIDASNVYGSTEHEARSIND LASHRGLLRQGIVQRSGKPLLPFATGPFPECN RDENESPPCFLAGDHRANEQLGLTSMHTLW FREINRIATELLKLNPHVDGDTTYFETRIXV AEIQHITYQHWLPKLIGEVGMRTLGEVHGYT ENGAGTMAFATAAFREFILLVRPLLIPGLI ENFQPIAQDHLPLHKAFFSPFRIVERGIDPLL RGLGFQVAGKMRYPSQLLTRELFSMAHT VALDLAANNQRGRDHGIPPYHDYRVCNLS AAHTFEDLKKMRYPSQLTERLFSMAHT VALDLAANNQRGRDHGIPPYHDYRVCNLS AAHTFEDLKNEIKVPEREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFFRLI DGDRLWYENPGVFSPAQLTQKTGTSLARILCI NADNITRVQBSQVFRVAEPPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKTRPRFIFVSGRQGEHISNISTS ANSTRSDASGATNDPGRVCSWEMGKTTDLR TOKKLESRUSTTEVDVAGGEHANNTTKW KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEERP 608 1958 A 4566 354 1135 FSFLCGVSGRIGLGDSEEDVYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGGHTIDLR TOKKLESRUSTTEVDVAGGVVIRVSCG LGVGMYKKLEGRUSTFTCSKCGG LGVGMYKKLEGRUSTFTCSKCGG GFDAW FSFLCGVSGRIGHEAYHEVPTTSFTLACQLS YKIRTTAPGKTHTAAIDERGRILTTGCKKCGG LGVGMYSKLEGRUSTSSCSGGTGTVSH HOSDICTSWPRPIEGSLHHVPDLSCRGWHTLL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW FSFLCGSVSGRIGHEAYHEVPTTSFTLAQLS VYRIRTAPGKTHTAAIDERGRILTTGCKKCGG GFDAW FFFETESSVAQAGVQWRDLGSLQAPPPGF FPSCLSLPSSVDYRPPLRFANFFYLTGTF HRSPBGLGLILTTSGCRNSGSGGTGTVSH RAPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHVCTLCLPSSSDSASAFRVARTT			Į.				•
GRPPPVIAWTKGGSQLSVDRRIFLUTSGTLR SGVALHDQGYECQAVINGSQKVVAHLTVV PRVTPVFASIPDITTVEVGANVQLPCSSQGEP EPAITWNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPI VSRNGDPFVATSIVESAIATVDRAINSTHLF DSRRSPNDLLALFRYPRDPTYTVEQARAGEII ERTLQLIQEHVQHGLMVDLNGTSYHYDDLV PQYLNLIANLSGCTAHRRYNGSDMCFHGK RIHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGALMPPRLVSTTLL GTETVTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDOQHCSNVCSNDPPCFSVMIPPI DSRARSGARGHFVRSSPVCGSGMTSLLMM VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHIGLLRGQFIVQRSGKPLLPFATGPPTECK RDENSSPPCFLAGDHRANEQLGITSMHTLW FREIBNATELLKLMPHWDGLYFLTFATGPPTECK RDENSSPPCFLAGDHRANEQLGITSMHTLW FREIBNATELLKLMPHWDGHTYYETRKV AEIQHITYQHWLPKLIGEGWARTLGEVHGYT PGNAGIFNAFATDAAFRGHTLYNPLLIPGLI ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL ROLFOVAGKMRVPSQLLINTELTERLFSMATH VALDLAANINGRGDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNEFBREKLRLVSGTLNID LFPALVVEDLYPGSRLGFITMCLISTGFRIL DGDRLWYENPGVFSPAQLTQIKGTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSLI FSYQEDKPTKTRPRRIFSVGQCGEHISNITS AFSTRSDASGNTDFQRVCSWEMQKTHIDLR TQIKKLESRLSTIECVDAGGESHANNTKWK KDACTICECKDGQVTCFPACAPPATCAVPNV PGACCPVCLQKRAEEKP FSPLCGWSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQGCDGTFILLTQSGKVLAGGENHKI GLQCMSGINHEAYHEUYSTLAKQLIS YKIRTIAPGKTHTAADERGRILTFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVIKVSGGD EFTIAATDDNHFAWGNGGGNGLAMIPTERF HGSDICTSWPRIPGSLHHVPDLSCRGWHTLL VEKKVLNSKTIRSNSSGLSIGFOSSSPGGGG GGPDAW FPFETESRSVAQAGVQWRDLGSLQAPPFGF FPSCLSIPSSWDYRRPLRFANFFVFILVETGF HGSDICTSWPRIPGSLHHVPDLSCRGWHTLL VEKKVLNSKTIRSNSSGLSIGTOSSSPGGGGG GGPDAW FPFETESRSVAQAGVQWRDLGSLQAPPFGF FPSCLSIPSSWDYRRPLRFANFFVFILVETGF HRSRDGLDLITX/GDPPASASQSAGTTGVSH RARPINLANNVYSFAVTYCLLYNISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECROVISANHCCTLCLPSSSDSASAFRVARTT			1				l
SGVALHDOGQYECQAVNIIGSOKVAHLTVV PRNTPYFASIPSDTTVEVGANVQLPCSSQGEP EPAITWNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMULSVNVPI VSRNGDPFVATSIVEALATVDRAINSITRTHLF DSRPRSPNDLLALFRYPRDPYTVGARAGEIE ERTLQLIGEHVQHGIMVDLNGTSYHYNDLV PQVLNLIANLSGCTAHRRVNNCSDMCFHQK RITHDGTCNNLQHEMWGASLTAFERLLKSVY ENGFNTFRGINPHRLYNGHALPMPRLVSTTLI GTETVTPDCFTHMLMQWGQFLDHDLDSTV VALSQARFSDGOHCSNVCSNDPFCFSVMIPP DSRARSGARCMFFVRSSPVCOSGMTSLLMN VYPREQINGLTSYIDASNVGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPFTECK REDENSPHPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIVG AEIGHITYQHWLPKILGEVGMTLIGEYHOVT PGINAGIFNAFATAAFRGHTLVNPLLLPGLI ENFQPIAQDHLPLHKAFFSPRINVFGSIDPHL RGLFGVGAGMRVPSQLLNTELTERLFSMAH VALDLAAINIQRGRDHGIPPYHDYRYCYOLS AAHTFEDLKMEKNPEIREKLRRLYGSTLND LFPALVYEDLYPGSRLGFTLMCLLSTOFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARLCI NADNITRVQSDVFRVAFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFGRRSLI FSYQEBKPTKKTRRRHSPSGRQGEHLSNSTS AFSTRSDASGTNDFQRVCSWEMGKTITDLR TQIKKLESRILSTTECYDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRIGHLSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLAGGLNEFNKI GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKRTIAPGKTHTAADDERGRLLTFCGNKCGG LQGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC LGYGNYKKRLGINLLGGPLGGKQVRVSCGC HTAATDDNHFAWGNGGNGRLAMTPTERP HGSDLCTSWPRPGFSLHHVPDLSCFGCWHTILL VKKVLNSKTIRSNSGLSLSIGTVPQSSSPGGGGG GGPDAW PFFFETESSRSVAQAGVQWRDLGSLQAPPPGF PFSCLSLPSSWDYRRPPLRPANFPTJLVETGC HRSREDGILLT/SGGDPAASSGSAGTGVSH RAPPRINLRNVYSFAVTYCLNYSLAMSSTL KLSPHVLSGS 609 1959 A 4567 1 467 ECRGVSSAHCCTLCLPSSSDSASAFRVARTT							
PRVTPVFASIPSDITTVEVGANVQLPCSSQGEP EPATIWNKDGVQVTESPKISPEGELTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPI VSRNGDPFVATSIVEALATVDRAINSTETHLE, DSRPRSPNDLLALFRYPRDPYTVEQARAGEII ERTLQLIQEHVQHGMVDLNGTSYHYNDLV PQVLNLIANLSGCTAHRVNCSDMCFHQK RTHQGTCNRLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLL GTETVTPDEQFTHMLMQWGGPLDHDLDSTV VALSQARFSDGQHCSNVCSNDPPCSVMIPPP DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIRQ LASHRGLLRQGIVQRSGKPLLPPATCPPTECK RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTITYSTERKIV AEIQHITYQHWLPKLIGEVGMRTLGEVHGYT PGINAGIFNAFATAAFRFGHTLVNPLLLFGAL ENFOPIAQOHLPLHKAFRFGHTLVNPLLLFGAL ENFOPIAQOHLPLHKAFRFGHTLVNPLLLFGAL FREHNRIATELLSNAPHDGVAGCHARAGEI ENFOPIAQOHLPLHKAFRGGTLVNPLLLFGAL FREHNRIATELLSNAPHDGVAGCHARAGEI BOBRLWYENFOVESPSFUNNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAH VALDLANIQRGRDHGIPPHDYTNCLLSTQFKRL DGDRLWYENFOVESPSFUNNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAH VALDLANIQRGROHEJPHDYTGSTLONG LPPALVYEDLVPGSRLGPTLMCLLSTQFKRL DGDRLWYENFOVESPSKLFRVAGCGEHENSITS AVSTRSDASGINDFQRVCSWERGHELSNSTS AVSTRSDASGINDFQRVCSWERGHELSNSTS AVSTRSDASGINDFQRVCSWERGKTITDLR TQUKKLESRINSTTGCVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEERP TGVKLESRINSTGCVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEERP GROWNSGINHEAYHEVPVTTSFTLAKQLST YKRTIAPGKTHTAADDERGRLLTFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVRVSCGG EFTIAATDDNIBFAWGNGGNGRLAMTPTERP HGSDCTSWPPPIFGSLHTQCNKCGG LGVGNYKKRLGINLLGGPLGGKQVRVSCGG EFTIAATDDNIBFAWGNGGNGRLAMTPTERP HGSDCTSWPPPIFGSLHTPOLSCCGWHTILL VEKVLNSKTIRSNSSGLSIGTVPQSSSPCGGGG GGPDAW FFFETESRSVAQAGVQWRDLGSLQAPPFGF FFSCLSLPSSWDYRRPPLRPANFFYFLVETGF HGSDCTSWPPPIFGSLDLTTNSGDPSASQSAGTGVSH RAPPRINLRNVYSFAVTYCLNYISLAMSSTL KLSPHVLSGS 609 1959 Å 4567 I 412 FFFFETESRSVAQAGVQWRDLGSLQAPPFGF FFSCLSLPSSWDYRRPPLRPANFFYFLVETGF HRSSDGLDLTTNSGDPSASQSAGTGTGVSH RAPPRINLRNVYSFAVTYCLNYISLAMSSTL KLSPHVLSGS 6010 1960 A 4570 697 467 ECGRYSAHCCTLCLPSSSDSASAFRVARTT	1		ĺ				
EPAITWNKDGVQVTESGKPHISPGEITTINDV GPADAGKPECVARNTISGASVSMVLSVNVPI VSRNGDPFVATSIVEAIATVDRAINSTRTHLF DSRPRSPNDLLALFRYPRDPYTVEQARAGEII ERTLQLIGEHVQHGLMVDLNGTSYHYNDLV PQVLNLIANLSGCTAHRRVNNCSDMCFHQK: RITHDGTCNLQHPMWGASLTAFERILKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTL GTETVTPDEGTTHMLMQWGGPLDHDLDSTV VALSQARFSDGQHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFFVRSSPVCSSGMTSLLMNS VYPREGNOLTSYIDASNVGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPTECK RDENESPIPCFLAGDHRANDGLGLTSMHTLW FREHNRIATELLKLNPHWGDITTYTETRINV AEIQHITYQHWLPKILGEVGMRTLGEYHGYV PGINAGIFINAFATAAFREHLVNPLLIPGLI ENFQPIAQDHLPLHKAFFSFENVEGGIDPLL RGLFGVAGKMRVPSQLLNTELTESMAHT VALDLAAINIQRGRDHGPYHDYRVYCNLS AAHTFEDLKNEKNPEIREKLRLYGSTLNID LEPALVVEDLVPGSRLQGPAFSYHFRGRSSLI DGDRLWYENPGVFSPAQLTQIKGTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVII LEPALVVEDLVPGSRLQGFNAFSYHFRGRSSLI FSYGEDKPTKXTRPRKIPSVGRQGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAPVNN KURACTICECKDGQVTCFVEACPPATCAPVNN KURACTICECKDGQVTCFVEACPPATCAPVNN GLNQCMSGIINHEAYHPYTTSFTLAKQLSF HGSDICTSMPPHFGSLHAGLTFGCNKCGG LOVGNYKKRLGINLLGGPLGKGVURVSCGE EFTIAATDDHIBFAWGNGGNGRLAMTPTERF HGSDICTSWPPHFGSLHYDLSCRGWHTILLVEXULNSTHIPSVGRQGGGGGDAW GGPDAW GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GOPPAN GROSS GOPPAN							
GPADAGRYECVÁRNTIGSASVSMULSVNIVPI VSRNGPPVATSIVEALTVRAINSTETHLF DSRPSPNDLLALFRYPRDPYTVEQARAGEII ERTLQLIQEHVQHGLMVDLNGTSYHYNDLV PQYLNLIANLSGCTAHRAVNNCSDMCFHQK RIHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHALYMALPMRALVSTTL GTETVTPDEQFTHMLMQWGOFLDHDLDSTV VALSQARFSDQHCSNVCSNDPFCSVMIPPI DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSOKPLLPPATGPPTECK RDENESPPOFLAGDHRANRQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTTYYETRKHL AEIGHITYQHWLPKLIGEVGMRTLGEYHGVT PGINAGIFNAFATAAFRFGHTLNYPLLPFGLL ENFQPIAQOHLPJHKAFFENIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAH VALDLAANIQRGRDHGYHDYHDYRVCYCILS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLOPTI-MCLLSTQFKCHL DGDRLWVENPGVFSPAGFHGYGSCDEIPRVI LRVWQDCCDCTRTGGDHSYHFRGRRSLI FSYQEDKPTKKTRPKKIPSVGRQGEHLSNSTS AFSTRSDASGTNDFQRVCSWEMQKTITDLR TQKKLESRUSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEERF GON GNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAFGKTHTAADDERGRLLFTGCNKCGG LGVGNYKKRLGINLLGGELGGKQVKVSCCG LGVGNYKKRLGINLLGGFLGGKQVKVSCCG LGVGNYKKRLGINLLGGFLGGKQVKVSCCG EFTIAATDDNHIFAWGNGGNGRLAMTPTERF HGSDICTSWPPPIFGSLHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPCGGGI GGFDAW 609 1959 Å 4567 I 412 FPFPETESRSVAQAGVQWRDLGSLQAPPPGFI PFSCLSLPSSWDYRRPPLRPANFFVLVETGF HRSSRDGLDLTJNSGDPASASQSAGTIGVSH RAPPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 Å 4570 697 467 ECRGVISAHNCCTLCLPSSSDSASAFRVARTT		1	'				EPAITWNKDGVQVTESGKFHISPEGFLTINDV
DSRRSSPNDLLALFRYPRDPYTYEQARAGEII ERTIQLIQEHVQHICMPLORIGSYHYNDLV PQYLNLIANLSGCTAHRRVNNCSDMCFHQK RITHDGTCNNLQHPMWGASLTAFERLLKSVY ENGRITPRGRIPPHILVSTIL GTETVTPDEQFTIMLMQWGQFLDHDLDSTY VALSQARFSDQQHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFFYRSSPVGSGMTSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLIRQGIVQRSGKPLLPFATGPFTECK RDENESPPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDDDTIYYETRIKVC AEQHITYQHWLPKILGEVGMRTLGEYHGYT PGINAGIFNAFATAAFRGHTLVNPLLLPGLL ENOPJIAQOHLPLHAFSPFRIVNEGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHI VALDLAANINGRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPERIEKLKRLYGSTLND LFPALVVEDLVPGSRLGPTLMCLLSTQFKRIFL DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRQQFNAFSYHFRGRRSLI FSVQEDKYTKKTRPISVGRQGEHLSNSTS AIFSTRSDASGTYNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTGVPEACPPATCAVPVN PGACCPVCLQKRAEEKP FSFLCGVSGRLGJDSEEDYYTPQKVDVPKAI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAFGKHTHAMDERGRLLTFGGNKCGG LGVGNYKKRLGFNLLGGFLGGKQVRVSCCG EFTIAATIDDNHIFAWGNGGNGLAMTPTERP HGSDICTSWPRPIEGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVPQSSSPGGGG GGPDAW 609 1959 A 4567 I 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRANFFVFLVETGF HRSRDGDLDLTTSGPSLAHVDSLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1950 A 4570 697 467 ECROVISAHNCCTLCLPSSSDSASAFRVARTT			[:				GPADAGRYECVÄRNTIGSASVSMVLSVNVPD
ERTLQLIQEHVQHGLMYDLNGTSYHYNDLV PQYLNLLANLSGCTAHRRVNNCSDMCFHQK RTHDGTCNNLQHPMWGASLTAFERLLKSVY ENGRNTPRGINPHRLYNGHALPMPRLVSTTLL GTETVTPDEQFTHMLMQWGGFLDHDLDSTY VALSQARFSDGQHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFFVRSSPVCGSGMTSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECK RDENESPPCTLAGHNAEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRILVC AEIQHITYQHWLPKILGEVGMRTLGEYHGYC AEIQHITYQHWLPKILGEVGMRTLGEYHGYC PGINAGIFNAATAARFGCHTLVNPLLLPGLL ENRQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRWPSGLJNTELTERLFSMAHT VALDLAANIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKKRYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLH DGDRLWYENPGVFSPAQLTQHQTSLARILCI NADNITRVQSDVFRVAFFPHGYGSCDEIPRNV LRVWQDCCEDCRTRGGFNAFSYHFERGRRSLI FSYQEBKPIKKTRPRRIPSVGRQGEHLSNIST AFSTRSDASGTNDPGRVCSWEMQKTITIDLR TQIKKLESRUSTTECVDAGGESHANNTKWK KDACTLCECKDGQVTCVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSPLCGVSGRGLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSS YKIRTLAGKHTHAADERGRLLTHFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHFAWGNIGGNGRLAMITFERP HGSDICTSWPRPIEGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRANFFVFLVETGF HFSRDGDLDLTXSGPASAGSAGTIGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHNCCTLCLPSSSDSASAFRVARTT	İ		•				VSRNGDPFVATSIVEAIATVDRAINSTRTHLF
PQYLNLIANI.SGCTAHRRVNNCSDMCFHQK: RIHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLYSTTIL GTETYTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDGOISNVCSNDPPCFSVMIPPI DSRARSGARCMFFVRSSPVCGSGMTSLLMNS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECK RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHRIATELLKLHPHWDGDITYYETRKIVG AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATAAAFRFGHTLVNPLLLPGLE ENFQPIAQDHLPLKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHI VALDLAANIQRGRDHGIPPYHDYRVYCNLS AAHTTEDLKNEIKHPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGFTLMCLLSTOFKRLI DGDRLWFENFPGSPAQLTOJKOFSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSLI FSYQEDKFTKKTRFRKIPSVGRQGEHLSNISTS AIFSTRSDASGNTNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP FSELCGVSGRLGLDESEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGINHEAYHEVYTTSFILAKQLST YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGG LGVGMYKKRLGINLLGGPLGGKQVIR VSCGD EFTIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWFRPIFGSLHHVPDLSCRGWHTLI VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLSPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDYRRPPLRFANFFVFLVETGF PFSCLSLSLPSSWDASAFRVARTT 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT		}					DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
RIHDGTCNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTL GTETVTPDEQFTHMLMQWGQFLDHDLDSTY VALSQARFSDGQHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFVRSSPVCGSGGSMSLLMN VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPPCFLAGLTSMHTLW FREHNRIATELLKLNPHWDGDTTYYETRKIVC AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATAAFRFGHTLVNPLLLPGIL ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRYOQLLNTELTERLFSMAHI VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTTEDLKNEIKNPEIRERLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTOFKRIF DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEPPHGYGSCDEIPRVI LRVWQDCCEDECTRIGQPNAFSYHFRGRRSIL FSYGEDKPTKKTRPRKIPSVGRQGEHLSNSTS AFSTRSDASGVINDFQRVCSWEMQKTITDLR TQIKKLESRUSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFYBACPPATCAVPVN PGACCPVCLQKRAFEKP 608 1958 A 4566 354 1135 FSFLCGVSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQCGCGTFLLTQSGKVLACGLNEFNKI GLNQCMSGINHEAVHEVPYTTSFLAKQLS YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFSMGNGGNGRLAMFTFERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVPQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPGFF PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRSRDGLDLLTYSGDPPASASQSAGTIGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
ENGFNTPRGINPHRLYNGHALPMPRLVSTTL GTETVTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDQOHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFFVRSSPVGSGMTSLLMNS VYPREQINQLTSVIDASNVYGSTEHEABSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPPCFLAGDHRANRQLGLTSMHTLW FREHINRIATELLKLNPHWDGDTIYYETRIKV AEIQHITYQHWLPKILGEVGMRTLGEYHGVE PGINAGIFNAFATNAFRFGHTLVNPLLLPGLE ENFQPIAQOHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTEELFSMAHI VALDLAANINQRBHGIPPYHDVRYVCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGFTLMCLLSTQFKRLI DGDRLWTENFPGYSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEPPHGYGSCDEIPRVI LRVWQDCCEDCTRTGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AIFSTRSDASGNINDFQRVCSWEMQKTITDLR TQIKKLESRUSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP FSFLCGVSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAADDERGRLLTFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTILATDDNHIBFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILV VEKVLNSKTIRSNSSGLSIGTVPQSSSPCGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDYRRPPLRRANFFVFLVETGF PRSCLSLPSSWDASAFRVARTT 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	ļ		ļ				
GTETYTPDEOFTHMLMQWGQFLDHDLDSTY VALSQARFSDQHCSNVCSNDPPCFSVMIPP DSRARSGARGMFYNSSPVCGSGMTSLLMM VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTTYYETRKIVQ AEIQHITYQHWLPKILGEVQMRTLGEYHGYT PGINAGIFNAFATAAFRFGHTLVNPLLLPGLI ENFQPIAQDHLPLHKAFFSFFRIVANEGGIDPLL RGLFGVAGKMYSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTUMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSLI FSYQEBKPTIKKTRPRKIPSVGRQGEHLSNISTS AFSTRSDASGTNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGDSEEDYYTPQKVDVPKAI IIVAVQCCCDGTFILLTGGKVLAGCLINEFNKI GLNQCMSGIINEAYHBEVPTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGFLGKQVIRVSCGD FFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVPQSSSPGGGGI GGPDAW 609 1959 A 4567 I 412 FFFPETESRSVAQAGVQWRDLGSLQAPPGFF PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRGDLDLLT/S/GDPPASASQSAGITGVSH RAPRRILRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHNCCTLCLPSSSDSASAFRVARTT							l
VALSQARFSDOQHCSNVCSNDPPCFSVMIPPI DSRARSGARCMFFVRSSPVCGSGMTSLLMMS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIV AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATAAFRFGHTLVNPLLLPGLI ENFQPIAQDHLPHKAFFSSFRIVNEGGIDPLI RGLFGVAGKMRVPSQLLNTELTERLSMAHI VALDLAANNIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKAFRKIPSVGRQGEHLSNSTS AVFSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI GLNQCMSGINHEAYHEVPYTTSTTLAKQLST YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVIKVSCGD LGVGNYKKRLGINLLGGPLGGKQVIKVSCGD LGVGNYKKRLGINLLGGPLGGKQVIKVSCGD LGVGNYKKRLGINLLGGPLGGKQVIKVSCGG LGVGNYKKRLGINLLGGPLGGKQVIKVSCGG GFDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWDLGSLQAPPFGF FFSCLSLPSSWDYRRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWDLGSLQAPPFGF FFSCLSLPSSWDYRRPLRPANFFVFLVETGF HRFSRGGLDLLT7/S/GDPPASASQSAGITGVSH RAPRRINLRNVTYSFAVTYCLNYISLAMSSTL KLSFFVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							l
DSRARSGARCMFFVRSSPVCGSGMTSLLMNS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLIRQGIVQRSGKPLLPPATGPPTECM RDENESPPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIVV AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATAAAFRFGHTLVNPLLLPGIL ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLI RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANNQRGRHGIPPYHBYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVGSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSLI FSYQEDKPTKKTRPKIPSVGRQEHLSNSTS AVFSTRSDASGNTNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRLGLDSEEDYYTPQKVDVPKAI GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAADERGRLLTFGCNKCGG LGVGNYKKRLGINILLGGPLGGKQVIRVSCGD EFTTIAATIDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFFTESRSVAQAGVQWRDLGSLQAPPPGF PFSCLSLPSSWDYRPLRPANFFVFLVETGF HRSRDGLDLLT/SIGDPPASASQSAGITGVSH RAPPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	ļ						l
VYPREQINQLTSYIDASNYYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCTLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTIYYETRKIVV AEIQHITYQHWLPKILGEVGMRTLGEYGHGYU AEIQHITYQHWLPKILGEVGMRTLGEYGHGYU PGGNAGIFNAFATNAAFRFGHTLVNPLLLPGLI ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLLI RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFRKIL DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRQGFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AFSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESRULSTTECVDAGGESHANNTKW KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDFFLLT\QSGKVLACGLNEFNKI GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGD LGYGNYKKRLGINLLGGFLGKQVIRVSCGD LGYGNYKKRLGINLLGGFLGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VERVINSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGQVQWRDLGSLQAPPPGF PFSCLSLPSSVDVRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAF\RVARTT							, · · · · · · · · · · · · · · · · · · ·
LASHRGLIRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCFLAGDHRANEQLGLTSMHTILW FRENRNIATELLKLNPHWDGDTTYYETRKIVG AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATTAAFRCHITLVNPLLLPGIL ENFQPIAQDHLPLHKAFFSPFRIVNEGGIPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANNQRGRHGIPPYHDYRVYCNLS. AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AVFSTRSDASGTINDFQRVCSWEMQKTITIDST AVFSTRSDASGTINDFQRVCSWEMQKTITIDST TQIKKLESRUSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEKP 608 1958 A 4566 354 1135 FSFLCGVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAVHEVPYTTSFTLAKQLSS YKIRTIAPGKTHTAAIDERGRLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKITRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDVRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	1	1					
RDENESPIPCFLAGDHRANEQLGLTSMITLW FREHNRIATELLKLNPHWDGDTIYYETRKIVC AEIQHITYQHWLPKILGEVGMRTLGEYGYT PGINAGIFNAFATAAFRFGHTLVNPLLLPGLI ENFQPIAQDHLPHKAFFSPFRIVNEGGIPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTTEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLISTOFKRIE DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVRVAEFPHGYGSCDEIPRVI LRVWQDCCDCCRTRGQFNAFSYHFGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AIFSTRSDASGATNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVBACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVYJSPAVTYCLNYISLAMSSTL KLSFNVLSGS 610 1960 A 4570 697 467 ECROVISAHCCTLCLPSSSDSASAFRVARTT	1	ľ					
FREINRIATELLKLNPHWDGDTIYYETRKIVG AEIQHITYQHWLRKILGEVGMRTLGEYHGYD PGINAGINARATINAFREGHTLVNPLLLFGLI ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRYSQLLNITELTERLFSMAHI VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTILMCLLSTOFKRLE DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRYQSDVFRVAEFPHGYGSCDEIPRVII LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPKIPSVGRQGEHLSNSTS AFSTRSDASGTNDFFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGG LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRFANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RAPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							, ,
AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATNAAFRGHTLVNPLLLPGIL ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLI RGLFGVAGKMRVPSQLLNTELTERLFSMAFTI VALDLAAINIQRGNDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKKKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGGFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AFSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDLCTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							l
PGINAGIFNAFATVAAFREGHTLVNPLLLPGLI ENFQPIAQDHILPIHKAFFSSFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLF DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGGPFNAFSYHFRGRSRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A/FSTRSDASGVINDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN FGACCPVCLQKRAEEKP FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFFLAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							
ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLI RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLF DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCKTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AIFSTRSDASGTNDFQRVCSWEMQKTITDLR TQIKKLESRILSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLCGVSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFITIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEK VLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGF PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLTVS/GDPPASASQSAGITGVSH RARPRINLRNVTYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							
RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRIGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVTYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAFRVARTT	ĺ	[ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLI DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS ANFSTRSDASGTNDFQRVCSWEMQKTITDLR TQIKKLESRLSTTECVDAGGESHANNTK WK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFILITQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLTS/GDPPASASQSAGITGVSH RARPPINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLE DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCTRTGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQEGHLSNSTS A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAF\RVARTT							
DGDRLWYENPGVFSPAQLTQIKQTSLARILCI NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS AFSTRSDASGTNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEK VLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	1						AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
NADNITRVQSDVFRVAEFPHGYGSCDEIPRVI LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A\FSTRSDASGTNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTK WK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 I 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVTYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAFRVARTT	İ						LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR
LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLI FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTK\WK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC\GVSGR\LGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD GFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTIL VEK\LNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGF1 PFSCLSLPS\WDYRPPLRPANFFVFLVETGF HRFSRDGLDLLT\S\GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT							DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS A\FSTRSDASG\TNDFQR\CS\WEMQKTITDLR T\QIKKLESR\LST\TEC\DAGGESHAN\NTK\WK KDACTICECKDGQ\TCF\VEACPPATCA\VP\N PGACCP\VCLQKRAEKP 608 1958 A 4566 354 1135 FSFLC\GV\SGRUGLD\SEED\YT\PQ\K\VD\VP\KAI II\VA\VQCGCDG\TFLLT\QSGK\VLACGL\NE\F\NK\II\VA\VQCGCDG\TFLLT\QSGK\VLACGL\NE\F\NK\II\VA\VQCG\TG\THAIDER\GR\LT\FGC\NK\CG\Q L\G\VG\NY\K\RL\G\IN\LG\GPL\GG\KQ\VIR\VS\CG\D E\FTI\AATDD\N\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	1						NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAF\RVARTT	{						LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
TQIKKLESR\LSTTECVDAGGESHANNTKWK KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT							
KDACTICECKDGQVTCFVEACPPATCAVPVN PGACCPVCLQKRAEEKP 608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	1				•		
PGACCPVCLQKRAEEKP	1				Į		
608 1958 A 4566 354 1135 FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAI IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKI GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATIDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT		[1			
iivavqcgcdgtflltqsgkvlacglnefnki glnqcmsgiinheayhevpyttsftlakqlsf ykirtiapgkthtaaidergrlltfgcnkcgq lgvgnykkrlginllggplggkqvirvscgd eftiaatddnhifawgnggggrlamtpterp Hgsdictswprpifgslhhvpdlscrgwhtill vekvlnsktirsnssglsigtvfqssspggggi ggpdaw 609 1959 A 4567 1 412 Ffffetesrsvaqagvqwrdlgslqapppgff pfsclslpsswdyrrpplrpanffvflvetgf Hrfsrdgldlltt/s/gdppasasqsagitgvsh Rarprinlrnviysfavtyclnyislamsstl klsfhvlsgs 610 1960 A 4570 697 467 Ecrgvisahccttlclpsssdsaafrvartt	608	1958	A	4566	354	1135	
GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMIPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT	***		1.	7500	334	.133	
YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGI GGPDAW 609 1959 A 4567 I 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT					.		GLNOCMSGIINHEAYHEVPYTTSETI AKOI SE
LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGG GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFRVARTT		. 1		ļ			
EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAF\RVARTT	1			1			
HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILL VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT					ļ		EFTIAATDDNHIFAWGNGGNGRLAMTPTERP
VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGI GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT	I						HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
GGPDAW 609 1959 A 4567 1 412 FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT	ł			ļ	ļ		VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT		<u> </u>			ĺ		
PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT	609	1959	Α .	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFNVARTT							
RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS 610 1960 A 4570 697 467 ECRGVISAHCCTLCLPSSSDSASAFNVARTT							HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
610 1960 A 4570 697 467 ECRGVISAH\CCTLCLPS\SSDSASAF\RVARTT				1	ĺ		RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1							
GTCDYAOLIFAFLVEMGFHHVGODGLHLL/N	610	1960	Α	4570	697	467	
			ľ	I			GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N
LVIRPPRPPKVLGLQA							LVIRPPRPPKVLGLQA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	.location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			l	peptide	504,000.00	/=possible nucleotide deletion, \=possible
ł	1	1		sequence		nucleotide insertion
611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
*	***	l ' '	1371		1370	WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
Ĺ						LCAATAVLLSAQGGPVQSKSPRFASWDEMN
[ĺ	i			ĺ	VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
						LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
					!	LOTOLKAONSRIOOLFHKVAQOORHLEKOHL
1	j	l			1	RIOHLOSOFGLLDHKHLDHEVAKPARRKRLP
1						EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
						QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR
						RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
		}				EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
{		i			ł	VHLGGEDTAYSLOLTAPVAGOLGATTVPPSG
ł						LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
i I					i	
					İ	GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
(10	1000		1635	160		WRGRYYPLQATTMLIQPMAAEAAS
612	1962	Α	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
613	1062	<u> </u>	1601	(00	201	GSPASASPVAGITGTRHHRTRG
613	1963	Α	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
j					j	SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
t		1				HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
						LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	Α	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
ì		(GGLFCAWVGTILLVVAMATDHWMQYRLSGS
1						FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
						GIFYMCAYRVHECRRLSTPR
615	1965	Α	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
						ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
						PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
						ERRPLPKNARPSPWVPALADEWNTLHQEVTT
						TRLPAGSQEPVKD
616	1966	Α	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
						SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
						AGLELPTSGDPPALASQSAGITGVTTVPSGPG
617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
						DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
						REKQLQELQQQQEEEERQRQQRREERRQQNL
						RARSREHPVVGHPDPALPPSGVNCSGCGAEL
						HCQDAR*
618	1968	Α	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
						SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
						MVGNETTYEDGHGSRKNITDLVEGAKKANG
í í						VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA
!						MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
1						HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV
						YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI
					•	AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
(l				j		AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
						CNPETFPSSNESRQCPNARCQFAFYGGESGYH
						RALLGLQIFNAFMFFWLANFVLALGQVTLAG
						AFASYYWALRKPDDLPAFPLFSAFGRALRYH
						TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
						KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
						IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
					İ	TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
						APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
						VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
						LNKTNKKAAES
619	1969	Α	4601	2	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY
لــــــــــــا		لــــت			J	THE PART OF THE PA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon. /=possible nucleotide deletion, \=possible nucleotide insertion
				coquento		GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN NQHVECNEICHRLSLTRPSMEKPCKS
620	1970		4606	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR KGHLEEEEDGEEGAETLAHFCPMELRGPEP LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF TGAFLLGYVAFRGSCQACGDSVLVVSEDVN YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL EDTIRQTSLRERVAGSAGMAALTQDIRAALS RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL VYAHYGRPEDLQDLRARGVDPVGRLLLVRV GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSMVSNGFR PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL IESVLKQVDSPNHSQTLYEQVVFTNIPSWD\ AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\ DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV LRHIGNLNEFSGDLKARGLTLQWYYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\ ALL\TWDACKGAANALSGDVWNIDNNF
621	1971	A	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\ NTLVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG
622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSVTVAYAIGMIL TFVVLGILMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI
	1974	A .	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH SLEENHFYSYPEEVDDDLICHICLQALLDPLD TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL QHCKKSSILVNKLLNKLLVTCPFREHCTQVL QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED CLSPGVHHCSEV
625	1975	Α	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP PPLLIPSS*LSP

SEO ID	SEQ ID	Met	SEQ	Predicted	Dandiet - J J	L Amino poid mouse of An-Alesia C. C. A.:
,					Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	I	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide		/=possible nucleotide deletion, \=possible
				sequence	1	nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI
020	1976	I A	4629	249	'	
	ł	l	ł		ł	LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP
			<u> </u>			ASASQVAGIAGTHH
627	1977	Α	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK
1		ľ				QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP
		l				PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP
			1		Ì	ARAYL
628	1978	Α	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR
1	.,,,	• •	10.0			TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK
			1			NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF
1	1	l	;		ŀ	
	1	1	i			YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK
1	l	1	1			VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF
						SYKSFAVIIFFVDNTRFFSFGF
629	1979	Α	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV
						KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH
1		l	!			PKLVFSOEGRYVKNTASASSWPVFSSAWNYF
		l	İ		1	AGWRNPOKTAFVERFOHLSCVLGKNVFTSG
	}	1	i			KHYWEVESRDSLEVAVGVCREDVMGITDRS
[[[
	[1			KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ
						EPALHRVGVYLDRGTGNVSFYSAVDGVHLH
	}					TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G
1	1					FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG
1	1			}		WSIFWVSLTVPFGICPLCASQEAVPWEVGLA
						NGDGTGNFPRRFWEIFL
630	1980	Α	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF
		l	[TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG
1		ł				FHDVGQDGLDLLTS*STPSASQSAEITGVSHC
						TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	
031	1201	A	40/4	933	014	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV
						AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ
1						NPVFLERRPRALHSSPGLTTQRILWAQGLWV
L			<u></u>			GAGSTGCSRGPRGEGVFREG
632	1982	Α	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP
	!					*LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP
1						ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1	1365	RTLGMEGERRASQAPSSGLPAGGANGESPGG
""		11	.570	1	1505	GAPFPGSSGSSALLOAEVLDLDEDEDDLEVFS
1 1			[KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP
1			ļ .			DLKDLFITVDEPESHVTTIETFITYRITKTSRG
		ì	1			EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII
						PPLPEKFIVKGMVERFNDDFIETRRKALHKFL
1	•					NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ
1 1	i]			GPGLLSRMGQTVRAVASSMRGVKNRPEEFM
			}			EMNNFIELFSQKINLIDKISQRIYKEEREYFDE
						MKEYGPIHILWSASEEDLVDTLKDVASCIDRC
			j	j		CKATEKRMSGLSEALLPVVHEYVLYSEMLM
						GVMKRRDQIQAELDSKVEVLTYKKADTDLL
						PEEIGKLEDKVECANNALKADWERWKQNM
						QNDIKLAFTDMAEENIHYYEQCLATWESFLT
			L l			SQTNLHLEEASEDKP
634	1984	Α	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ
						WISKAVYKHREMCGLTSTGRKSHGLEKDRM
[Ī	-			FPHAIGGSCRAA*RRKTLQFPCYH
635	1985	Α	4709	42	341	YIKQPDAKERRRTVHWKKETESEASEITIPPST
""	1703	^	7107	76	J-41	
						PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL
						WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT
						SED
636	1986	Α	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS
						RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I-Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence		{	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		İ	İ		j	ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
						LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
İ		ŀ				HAG*AGLELLTSGDPPASASRSAGITGVSHHA
		}		i		RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
		1				TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
						YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
ĺ						LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
į .		1]	WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	Α	4743	1040	699	QGLTLLPRMECSATITATICSLELPGSIDLPTSA
						S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
	i	ł				AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
"	.,,,	' '		527	-	WIRRRPCLPSGCLKMNREIGPLOHSLCCPGWS
						QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA
1	ĺ	[MARSRLTATSASQVQAILLPQPPGTTDSCSPS
	ŀ					PDHEQQPLSWVLPPPQKDMNPREQQVALGP QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
• • •	->>-	J	1,00			LQLAASPYFSPSWAECPQPVPAGTHATWCLA
	•					RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
		1				FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
642	1992	A	4798	1	487	QRDTDTGVHTGSGTHTHAHTPPEK GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
0.2	1772	11	.,,,,	1	,,,,	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
						TWWFGVKFAAGGLGTFHALLNTAVHVVMY
	ĺ	ļ				SYYGLSALGPAYQKYLWWKKYLTSLQLVQF
	1	ĺ				VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
}						QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
Ì] .				MVYFVGENNGDSSHNPVLAATGVGLDVAQS
						WEKAIRQALMPVTPQASVCTSPGQGKDHSK Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
	,					LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
		1				AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
645	1005		4905	450	126	SYKDIWGWPCLCGVLHAYIPLLV
645	1995	Α	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
		ľ				PLLAGLVAADAVASLLIVGAVFLCARPRRSP
						AQEDGKVYINMPGRG
646	1996	Α	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
						LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT HKQAVQCLKGPGQVARLVLERRVPRSTQQC
						PSANDSMGDERTAVSLVTALPGRPSSCVSVT
						DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
						KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE
						WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
						YYPAAVEVLHILRGAPQEVTLLLCRPPPGAL PELEQEWQTPELSADKEFTRATCTDSCTSPIL
						GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
L						EGTMGAKTERDLGPVP
647	1997	Α	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IVMPTYDLTDSVLETMGRVSLDMMSVQANT GPPWESKNSTAVWRGRDSRKERLELVKLSRK HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
					ļ	FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK LKWAKDHDEEAKKIAKAGQEFARNNLMGD DIFCYYFQTFPRNMPIYK
648	1998	A	4867	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTQ SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI LTEQHSKRVAVILNEFGEGSALEKSLAVSQG GELYEEWLELRNGCLCCSVKDNGLRAIENLM QKKGKFDYILLETTGLADPGAVASMFWVDA ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI NEATRQVALADAILINKTDLVPEEDVKKLRT TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV SWKDDTERTNRLVLLGRNLDKDILKQLFIAT VTETEKQWTTHFKEDQVCT
649	1999	Α	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV GQAGLELRTSGDPPASASQSAGITGVSHLA*P TSMPLLPFQRLCVYI
650	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR CPASFYLFLKYYLEAKFCA*GECAPSAGVGA GYKRGHKSCLLINCVVQI
651	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR PHMEPKASCPAAAPLMERKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL RRWADLLLVAPLDANTLGKVASGICDNLLTC VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR GQSPIPSRASSPSCSWAQVPGVALARCAGVC KPGDSWRVAACISGRCCSRGRRRGSGPRNPE QSFRGAWGPSFWGSWKSQRELSAGGAQAWP LLGSAGSGLRGEA
653	2003	Α	4965	2	283	FFFFI*DGVSLCHPGWNAVARSWLTATSASR VQAVSCFRLPSSWDYRHATMPG*FF*YF**R WGFT!LAILVLNS*PQVICPPWPPKVLTLQA
654	2004	A	4968	3 .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD IPAVGLGALGVIPPVKVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP PPCLTHLAAASCVVVWCGRWKRDSAECQCD HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

NO. of orling to be ordered to the country of the	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	1		hod				
Sequence							
1914 ng no first amino acid residue of peptide peptide peptide peptide peptide sequence 1914 ng no first amino acid of peptide sequence 1915 ng no first amino acid peptide	1		ł				
mino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence pertide s		uence					
Peptide Sequence	dence			914			
Peptide Sequence							
]		ļ	١.		Sequence	
				ļ ·		•	
AIWWEQKRQWILQTHWTLDKYGILADARLF							VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
658 2008 A 5017 I 292 FFFFKETESHSYTQAGVQWHDLQSLQPPPG						}	
							FGPQHRPVILRLPNRRALRLX*
HHVAQAGKLILTL*SANIGLSTSLPPIFLILE 559 2009 A 5018 17 338 RCHGGKSLTGGTPONWOGGLIVSEDWSHILE T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL MAACWAVHYKTHIMRPGLAVLPRLVLNSWS AIILUMPPKALGLQA 660 2010 A 5028 2 310 SRVDDFYGERGGCDECLCGHRGLRAVPLG HEGHLCLQPFGGF4*FLDVCRGCCPHPYDGST AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH HTGARWNH RQSCSSTQAKVQWFHTGPTVLCVCSFWIYQRGEPH HTGARWNH RQSCSSTQAKVQWFHTGPLQSQPPGLKQSSQ LSLPISDLRHKPPPLAIFSRAFTGSPYFAQAS LSLLGSSHPPTSASQSARITGVSHRAWPLK*F NINQYQTLTMN RQSCSSTQAKVQWFHTGPLQSQPPGLKQSSQ LSLPISDLRHKPPLAIFSRAFTGSPYFAQAS LSLLGSSHPPTSASQSARITGVSHRAWPLK*F NINQYQTLTMN THENACSQGSPSCAELLEVGAQAQLESCL) FARSQGRILALRTLLSQGTVVNAVTLIDHTY LHEACLGDHVACATTLLSAGGTVNANTTIDTOY TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVDQEIPHSQ TTENACSQGSPSCAELLEVGAQAQLESCL) SPITHEGASKGHHECDLILISWGIDVANATITIDGY TTENACSQGSPSCAELLEVGAQACAUSSISTLANG SPITHEMPHSSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHACAGGTTSTATAGAGTTAGASAACASLAPH SPITHAMPHACAGGTTSTATAGASAACASLAPH SPITHAMPHA	658	2008	Α	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
TYNSIVSPULGKWSPCLOGPGLSAVHTWPWL MAACWAVHYKTHMRPGLAVLPRLVLNSWS AIILLWPFKALGLQA A S028 2 310 SRVDDFVGERRGGCDECCGHRGLRAVPLG HPGHLCLQPPGGA*LDVCRGCCPHPVPGST AGSCRQKKTTPGPTVLCVCSFWIYQRGEPH HRTGARWNH HRTGARWNH HRTGARWNH HRTGARWNH RSVSSSTQAKVQWFHYGPLQSQPFGLKQSSQ LSLVSSRJPRTSAVGSARITGVSHRAWPLK*F NINOYOTITIMN LSPINSRJPHRAVPPRLAIFSFATGSPYFAQAS LELGSSIPPTSASQSARITGVSHRAWPLK*F NINOYOTITIMN LSPINSRJPHRAVPPRLAIFSFATGSPYFAQAS LELGSSIPPTSASQSARITGVSHRAWPLK*F NINOYOTITIMN LSPINSRJPHAVACARTLLEAGANVNAITIDGV TIPLFNACSQGSPSCABLLLEYGAQQALGSCLP SPINEGASGGHALARTILLSQGYNVNAVTLDHVTP LHEACLGPHVACARTLLEAGANVNAITIDGV TIPLFNACSQGSPSCABLLLEYGAQQALGSCLP SPINEGASGGHALARTILLSQGYNVNAVTLDHVTP LHEACLGPHVACARTLLEAGANVNAITIDGV TIPLFNACSQGSPSCABLLLEYGAQQALGSCLP SPINEGASGGHAGECDILISWGIDVDQEIPHSG TIPLYVACMAQQFHCIWNLIYAGAGYWRGKY WDTPLPGAGHGECDILISWGIDVDQEIPHSG TIPLYVACMAQQFHCIWNLIYAGAGYWRGKY WDTPLPGAGHGECDILISMGIDVVDQEIPHSG TIPLYVACMAQVAHAVASSSLAVENDO TIPLFNACSQGSPSCABLLLEYGAQQALGSCLP SPINEGASGGANVAAMAVSSSLAVENDO TIPLFNACSQGSPSAEFCKHAPAPPPSCA GPAEPSTTIFTQLATMAAFPILVHAELLPSSF WLRGGLGVVAAMAVAISSSLAVENDO TIPLFNACSQGSPSAEFCKHAPAPPPSCA GPAEPSTTIFTQLATMAAFPILVHAELLPSSF WLRGGLGVVAAMAVAISSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSTLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACSQGAMAAVASSSLAVENDO TIPLFNACS	(50	2000	<u> </u>	5010	-10	220	
MAACWAYHYKTHMRPGLAVLPRLVLNSWS	639	2009	A	2018	17	338	
*AIILLWPPKALGLQA * * * * * * * * *		j	l				
SRYDDFVGERRGGCDECLCGHRGLRAPPLG			1				
HPGHILCLOPPGGPA+FLDYCRGCCPHPVGST AGSCPROKKTTPGPTVLCVCSFWIYQRGEPH HRTGARWNH	660	2010	Α	5028	2	310	
AGSCPRQKKTTPQTYLCVCSFWIYQRGEPH HRTGARWNH	000	2010	l ''	3020	-	370	
HRTGARWNH	ļ		}				
LSI_PNSRDHRHVPPRLAIFSFAETGSPYFAQAS LELLGSSHPPTSAQSARITGSPYFAQAS LELLGSSHPPTSAQSARITGSPHAWPIK*F			١.	•			
LELLGSSHPPTSASQSARITGVSHRAWPLR*F	661	2011	Α	5050	752	431	
NLNQVQTLTMN				Ì			
FAASQGRLALRTILSQGYNVNAVTLDHVTP					120	100	
LHEACL GDHY ACARTILEAGANVNAITIDGY TPLFNACSQGSPSCAELLEYGAQAQ LESCLP SPTHEGASKGHHECLDILISWGIDVDQEIPHSG TPLYYACMAQQFHCIWNIJYAGAGVRKGKY WDTPLPGAGHIGSTOKLE*LFAMVEIWQ WDTPLPGAGHIGSTOKLE*LFAMVEIWQ WDTPLPGAGHIGSTOKLE*LFAMVEIWQ WDTPLPGAGHIGSTOKLE*LFAMVEIWQ WDTPLPGAGHIGSTOKLE*LFAMVEIWQ WRISP\$SFAHCASVYKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPPSCA GPAEPSTHIFTQLAMAPPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ CPAEPSTHIFTQLAMAPPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ CPAEPSTHIFTQLAMAPPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ QLFVIFLLLYLFTLGTNAIHISTIVLDRALHTP MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFIGCAJQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTYSLVTTISLVFHLPFHSSNQHE GFTYSLVTTISLVFHLPFHSSNQHE QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*1QRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHFAEEPPTLKGKRTVDVT KRTFATFGKDVSTREBRENGTLAMAPROPEPTLAMA TISPIPSTA TISP	062	2012	A	5054	48	103	
TPLFNACSQGSPSCAELLLEYGAQAQLESCLP SPTHEGASKGHHECDILISWGIDVDQEIPHSG SPTHEGASKGHHECDILISWGIDVDQEIPHSG TPLYVACMAQQFCLGWNLIYAGAGVRKGKY WDTPLPGAGHQSTQKLE-IFAMVEIWQ							
SPTHEGASKGHHECLDILISWGIDVIDEIPHSG							
TPLYVACMAQQFHCIWNLIYAGAGVRKGKY WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTGKHRLPAPVPFSCA GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSTSLTVRVLVKSQ GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSTSLTVRVLVKSQ QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE GFTVSLVTTSLVFHLPFHSSNQHE QQYSHNTGSAGHHAHCQVGHSPHVHYPSGC QQYSHNTGSAGHHAHCQVGHSPHVHYPSGC PL*IQRGLPSFNSLEGHSLKDSGHESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCQPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PIKPSVSYEUVDPGITARRC GFTVSLVTTSLVFHLPFDFQKTGPPL VGFKAQLSGLQLQPCLYKRR MILLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGFKAQLSGLQLQPCLYKRR GFTVSPLGTHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND GGPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSIFPPLLCGSQCLSEPVFGSHGPP RGCQHEAAPCPRGFGSDGLHHASAACASLPP SPILPVLLPELGPL.							
WDTPLPGAGHQSTOKLE*LFAMVEIWQ		i					
ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ	J						
GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ	663	2013	Α	5066	951	580	
WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ							
Colimor Coli							
		0014		5071	550		WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE 665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTITADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGROVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLTTLLAHVSPEPAGPSCDSLAOPG ASGV*VQHSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	004	2014	A	30/1	330	1	
TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE 665 2015 A 5074 496 692 QQYENTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRIL.GHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVYPQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 666 2016 A 5080 408 248 IMLESTSS*VYPQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPITTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLITLLAHVSPEPAGPSCDSLAQPG ASGV*VQHSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE	ľ						
GFTVSLVTTSLVFHLPFHSSNQHE G65 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATTGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							GFTVSLVTTSLVFHLPFHSSNQHE
HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGGSDGLHHASAACASLPP SPILPVLLPELGPL	665	2015	Α	5074	496	692	
HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATTGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYSIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTTPPLLTLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMILSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
666 2016 A 5080 408 248 IMILSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	666	2016	A	5080	408	248	
667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	667	2017	Α	5081	129	247	
RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	668	2018	Α	5086	852	233	
TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
GHYFNDTLVFLPPVGSEND 669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
669 2019 A 5101 I 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							•
ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	669	2019		5101		329	
RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL			·	J. J.	·		
SPILPVLLPELGPL					ļ	·]	
670 2020 A 5102 3 547 DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP							SPILPVLLPELGPL
	670	2020	Α	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI
						VGGTDQQYVSNDSGIYVSRIKENGAAALDGR LQEGDKILSVNGQDLKNLLHQDAVDLFRNA GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI FMVLVPVFALTMVAAWAFMRYRQQL
671	2021	Α	5105	672	400	RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF VLLLLLISLLCLYWKARKLSTLRSNTRKEKA LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA NHFVEVT
674		Ā	5153	3	2953	LTEDOPFDILQKSLQEANITEQTLAEEAYLDA SIGSSQQFAQAQLHPSSSASFTQASNVSNYSG QTLQPIGVTHVPVGASFASNTVGVQHGFMQH VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS MMTINNLDGSQIILKGSGQQAPSNVSGGLLV HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM GQQNTYNVNNLGIQQHHVQQGISFASASSPQ GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ TFAASGSPVIANHASPQLVGGQMPLQQASPT VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE MVMIDRMFNQEERASLSRDKRLALVDPEGFQ ADFCCSFKLDKAAHETQFGRSDQHGSKASSS LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS RSDHGTESKLSSILADSHLEMTCNNSFQDKSL RNSPKNEVLHTDIMKGSGEPQPDLQTKSLET TFKNILELKKAGRQPQSDPTVSGSVELDFPNF SPMASQENCLEKFIPDHSEGVVETDSILEAAV NSILEC
675	2025		5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR LYDEVQEVVYFPAVVHDNLGERLKCTYIEID QVPETYAVVLSRPAWLWGAEMGANEHGVCI GNEAVWGREEVCDEEALLGMDLVRLGLERA DTAEKALNVIVDLLEKYGQGGNCTEGRMVF SYHNSFLIADRNEAWILETAGKYWAAEKVQE GVRNISNQLSITTKIAREHPDMRNYAKRKGW WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE GYKLLNKHKGNITFETMMEILRDKPSGINME GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
676	2026	A	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG FTLLARMVSIS*PHDPPASASQSAGITGVSHRA RPT
677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR LNKRSFFMISPTDQQVHCWAWLKKIIMPKDS NLLLEDVTWKYTALNLIGPRAVDVLSELSYA PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
678	2028	A	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ GRIAKMPVKWIAIESLADRVYTSKSDVWAFG VTMWEIATRGMTPYPGVQNHEMYDYLLHG HRLKQPEDCLDELCKI**SPQSP
679	2029	A	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS
680	2030	A	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG
681	2031	Α	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK QSESAI
684	2034	Ā	5220	1	194	NLMKEMONLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS
685	2035	Α .	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide scq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FSSTMSLAKLLQERGISAKVYHSPISENPLQPL PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN FLASRPAETFLQEMYGLRPSRNPPDVGQLKM NLVDRLKRLGIARVVKNPGAQENGRCQEAEI GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM GSFAAPVCTSSPKMGVLKED
689	2039	A	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS GAPAGARGGPAKANSNPFEVKVNRQKFQILG RKTRHDVGLPGVSRARALRKRTQTLLKEYKE RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG GLLHKKTQQEGEREKPKSRKELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIQTLL SHKTPKSENRDKKEKPKPDAYDMMVRELGF EMKAQPSNRMKTEAELAKEEQEHLRKLEAE RLRRMLGKDEDENVKKPKHMSADDLNDGFV LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA SDPESNEEEGDSSGGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMEEQ LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPIKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKKK
690	2040	A	5261		304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK
691	2041		5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV
692	2042	A	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE IINTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
693	2043	A	5301	362	507	EVKDDDLFHSYTTIMALIHLGSSK EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
694	2044	A	5310	1	204	ACFPTNIVTLCHSIA RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA
695	2045	A	5315	125	1596	ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
i '	1	1				TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP
	ĺ					SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA
	ŀ					LATYYGSLFKLTDLKSLCSRGMYYGRDVNV
		l	<u> </u>			CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR
		ŀ				ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS
1					ļ	NVYITPAGSQGLPPHYDDVEVFILQLEGEKH
1	ľ	i	1			WRLYHPTVPLAREYSVEAEERIGRPVHEFML
		ĺ			11.0	KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST YQNNSWGDFLLDTISGLVFDTAKEDVELRTG
1	İ					IPROLLLOVESTIVATRRLSGFLRTLADRLEG
	ł	l				TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP
						GGKLPRLDSVVRLOFKDHIVLTVLPDODOSD
		l				ETOEKMVYIYHSLKNSRETHMMGNEEETEFH
						GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
		Ì				EKESLVLSLWTECLIQVV
696	2046	A	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSOGTIET
333	20.0				'	SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
			,			VEKLQVLLNCMTEIYYQFKKDKAERRLAYN
1 1						EEOIHKFDKOKLYYHATKAMTHFTDECVKK
						YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
						EEEVSKYQEYTNELQETLPQKMFTASSGIKHT
						MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE
						LAENNHILESGGSLTMDGGLRNVDCL
697	2047	A	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP
1						PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG
						VSPSWPGWSRTPDFR
698	2048	Α	5324	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC
						LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP
1						VGGLLMAFQKYSGETVQERKQKDRKALHEL
						KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA
600	0040		5224	(00	0.55	KKIEALLNLPRNPSVIDKQDKD
699	2049	A	5334	699	277	RPHGHLVCISSSAGLSGVNGLADYCASKFAA
ì						FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM
						FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
						YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK
700	2050	A	5344	3	614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE
/00	2030	^	2344	,	014	QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
						QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
1						PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
ł .						VETIQAQLLSTHDQPSVQALADEKNGAQTRP
						AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
						LLATNGTPL
701	2051	Α	5346	3	1383	HASVLFCRVMAASKTQGAVARMQEDRDGSC
					·	STVGGVGYGDSKDCILEPLSLPESPGGTTTLE
						GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV
						IADVKLVADFQRYILYWRKRFTEQPITDFCSV
]	IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
]						KQRLREILEQQQQERNDTNFHGVCMFCNEEF
		[LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL
						CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR
			. [KKQHRKINPKNREYDRFYVINYLELGKSWEE
						VQLEDDRELLDHQEDDWSDWEEHPASAVCL
						FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG
						LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS
]			. [KADLRTHMEETKHTSLLPDRKTWDQLEYYFP
						TYENDTLLWTLSDSESDLTAQEQNENVPIISE
702	2062		6357	2500	1540	DTSKLYALKQSSILNQLLL
702	2052	Α	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA
						CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRHIIVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK TVASRIAFFL
703	2053	Α	5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MVTLFQMWVVPLYFTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCPYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	Α	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI S
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	A	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	Α	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFIIPPV SAEIIRKMQQ
711	2061	Α	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749 ·	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KAPELLQGQSEDEQPDASQMHVYSLGMTLY WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR
713	2063	Ā	5506	22	478	RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPLHYMVIMN PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR CGHHEVDHFLCEMPALIRMACISTV
714	2064	Α	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR RLTSSVSCALDEAAAALTRMRAESTANAGQS DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM GRTALFHHSGGSSGYESLRRDSEATGSASSAP DSMSESGAASPGARTRSLKSPKKRATGLQRR RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG LACVSTRLRLAERRQQRLREVQAKHKHLCEE LAETQGRLMLEPGRWLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVAASAAIPGPQEVDV
716	2066	Α	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD SSSSERTMDLVLEMCNTNSIHWCGISGRQLG KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWOEDLDNMYLDTPRYRG RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	Α	5586	311	88	AVLKNMAPMTALGLLDLHILNLIIFLSAGEDF TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA
719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC
720	2070		5628	798	148	LPPAQIPEAWLLIANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN
722	2072		5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGHQGSCLAQCPSGFT RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
dence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
						AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
	•					PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRON
						KAEINPRTNGDRAACQTRTLRFVSNVTEADRI
İ						LLRWERYEPLEARDLLSFIVYYKESPFQNATE
						HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
						ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS- PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
						KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
		:				YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD
					-	CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
						NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL
						RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY RIDIHACNHAAHTVGCSAATFVFARTMPHRE
						ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
						ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
						HLALLPPGNYSARVRATSLAGNGSWTDSVAF
						YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
						LGFFYGKKRNRTLYASVNPEYFSASDMYVPD EWEVPREOISIIRELGOGSFGMVYEGLARGLE
						AGEESTPVALKTVNELASPRECIEFLKEASVM
						KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
						GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
						AGEIADGMAYLAANKFVHRDLAARNCMVSQ
						DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
						VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
						PLOLOELMSRCWQPNPRLRPSFTHILDSIQEEL
1						RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
						TPRDCSPQNGGPGH
723	2073	Α	5672	1	216	LAWIDNILPEKEKKETDKKRKKKGAHEDCD
						EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR SSSWSYSL
724	2074	A	5704	4235	940	ARGRRSRPVWAASWGGRGRPAARRRPRGLA
						ATMGFELDRFDGDVDPDLKCALCHKVLEDP
						LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
						LSAKELNHVLPLKRLILKLDIKCAYATRGCGR
						VVKLQQLPEHLERCDFAPARCRHAGCGQVLL RRDVEAHMRDACDARPVGRCQEGCGLPLTH
						GEQRAGGHCCARALRAHNGALQARLGALHK
						ALKKEALRAGKREKSLVAQLAAAQLELQMT
						ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE
						ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN
						GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT
						PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
						KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
						GDIHQEMDREELELEEVDLYRMNSQDKLGLT
						VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG
						DRIIQINGIEVQNREEAVALLTSEENKNFSLLI ARAELOLDEGWMDDDRNDFLDDLHMDMLE
						EQHHQAMQFTASVLQQKKHDEDGGTTDTAT
						ILSNQHEKDSGVGRTDESTRNDESSEQENNG
						DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
						NESFISADCTDADYLGIPVDECERFRELLELK
			٠			CQVKSATPYGLYYPSGPLDAGKSDPESVDKE LELLNEELRSIELECLSIVRAHKMQQLKEQYR
						ESWMLHNSGFRNYNTSIDVRRHELSDITELPE
						KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR
						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AAEGISCPSSEGAVGTTEAYGPASKNLLSITE DPEVGTPTYSPSLKELDPNQPLESKERRASDG
						SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR DRLLRERALKIREERSGMTTDDDAVSEMKM GRYWSKEERKQHLVKAKEQRRREFMMQSR LDCLKEQQAADDRKEMNILELSHKKMMKKR NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF LSVTTV
	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP WNTDRCFSNYSMVNTTNMTSAVVEFWERN MHQMTDGLDKPGQIRWPLAITLAIAWILVYF CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI
726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	A	5716	3	274	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE
728	2078	A	5737	1899	649	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD
729	2079	Α	5741		5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAAASPPPFQTPPRCPVPLL LLLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

AGRGLLYSRU-PARERLAYERFOOS ARAPALCARPRADVRAA(RAARTACFV APDVVAVLDSVVGGTOPACERKLNIOLOP) LDCGAAHLQHEISLQPLATFYPFRAGIT AVASV'NNYTAVEGPVHHVAQEDADSGY UNSRUVTAVGEPVHHVAQEDADSGY LMTSRIQMARVKVAACN'HSTEGDEVGAA AVGGWCALETRCTLQQDCTNISSQGHFWTS SEGPSRCAMTVLYSEDWYRGPYGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAFGI AVASV'NNYTAVEGPVHHVAQEDADSGYY SEGPSRCAMTVLYSEDWYRGPYGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAFGI AVCGWCALETRCTLQQDCTNISSQGHFWTS SEGPSRCAMTVLYSEDWYRGPYGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAFGI AVCGWCALETRCTLQQDCTNISSQGHFWTS GRINIKANETTYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCCNSQSRCEASPPTSPC CPRTILISPLAPVPTGGSQNILVPLANTAFFG CPRTILISPLAPVPTGGSQNILVPLANTAFFG CPRTILISPLAPVPTGGSQNILVPLANTAFFG CPRTILISPLAPVPTGGSQNILVPLANTAFFG CPRTILISPLAPVPTGGSQNILVPLANTAFFG CPRTILISPLAPVPTGGSQNILVPLANTAFTG CREATER WWYNESVYRCDQVY TTRKSQVFPLSLQLKGRPARFLDSPEPMTV VNCAMGSPDCSCQLGREDGHCLMWSD RLRGPLSVHFUVDTGPKAGGTRI HONDLHVGSBLQVLVNDTDCTELMRTOT ACTIMEGALPAPVPLVCWFERRGCVHONL WYMQNYSMAVHHGREPTLCXVLNSTLITCG GALSNASAPVDFTNGRAYADEVLK WYMQNYSMAVHHGREPTLCXVLNSTLITCG GALSNASAPVDFTNGRAYADEVLK HPGEPLTLVHIVSTKGAGKEQDSS,CQSHE RVKIGQVSCDIQIVSDRIHICSVNESLGAAV LPHTQVONFNQTIATLQLGGSETAINVSTVC LLLLSVVALFVFCTKSRRAERYWQKTLLQ) EMESSGRERKKGPABLQTDMTDLTKELNR GIPPLEYKHFVTRTFPKCSSLYERYYLSS LNSGGSQAQGETHLLGEWYTSIKKELLVDLI ASAARNPKLMLRYTSTEKELLVDLI ASAARNPKLMLRYTTSKELLVDLI ASAARNPKLMLRYTTSKELLVDLI ASAARNPKLMLRYTTSKELLDDTS VEDGRKKLNTLAHYKDFGGASLAMSLIDKK NILGRVRPLJAKSTGSTLGKFLLDDTS VEDGRKKLNTLAHYKDFGGASLAMSLIDKK NILGRVRPLJAKSTGSTLGKFLLDDTS VEDGRKKLNTLAHYKDFGGASLAMSLIDKK NILGRVRPLJAKSTGSTLGKFLLDDTS VEDGRKKLNTLAHYKDFGGASLAMSLIDKK RRTTMEEWLLREMBARRRNLWSPGG MOSLSVRAMDTDTLTQVKERLEAPCKNW SQWPRAEDVDLEWFASSTGSYLLREDAPCKAS RQSHRKKVLPETYLTRLISTKGTLQKFLDDD VEDGRKKLNTLAHYKDFGGABCRG PDTLHWKTNSLPLRFWVNILKNPGFVFDID TOHIDACLSVILAGAPGACSISDLQCKGBP KLLYAKEBETRKCVQRYYKGQIODMTPLSE RCSLASLITALHORD RATERTOR RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPHANSTGGACH RATTTREBENDERPH							
nucleotide sequence USSN 09496 uence unce corresponding of the properties of the popular of the					Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence SSAP			hod	ID NO:		nucleotide	
sequence peptide greided greid	nucl-	peptide			nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
sequence peptide greided greid	cotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence 914 min o first amino said residue of peptide peptide residue of peptide sequence Peptide seque		_	1	1	,		
amino acid residue of sequence peptide sequence peptide sequence of sequence of sequence sequ	•		l				
Peptide of peptide Pep	401.00		ł	1			
peptide			i	į .			
sequence nucleoide insertion				ļ		sequence	
FELMPSIDDNILKIKQGAKEQHELGFVSAR_ SDPPFQAQSYANLALNSEARAGRESQAR_ LARICLPHOAGGDAKKLTESYIQLGLQCA AGRGDLYSRLVSYPARERLFAYPERPQOS ARAAPAALCAFRADVRAAIRAARTACTV APDVVAVLDSVVQGTGPACERKLNIQLQP LDCCAAAHQPHESILQPLATPYPRAPGLT AVASVNSYTAVFLGTVNGRLKININESM VVSRRVVTVAVQEPYHHVMQEDPADSGYI LMTSHQMARVKVAACNVHSTCGDCVGAA AVCGWCALETRCTQQDCTNSSQGHFWT SEGFSRCPAMTVLPSEDVRQFYPGMIQLSI LPSLSGMEMACDYGNNIRTVARPQPAFGI IAYCNLLPRDQFPPFPPNQDHVTVEMSVRV GRNIVKANFTIYDGSSTIQVYPHATACTSCL QWPCFWCSQQHSCVSNQSRCEASPPITSPF CPRTLISPLAPVTIGSSGNILVYLANTAFFQ AALECSFGLEBFFAWWVBSVVRCDQVVI TIRKSQVFFLSQLKGRPAFLLSPEPPHTVI VVNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLOPMAGTCAPPERRAERFSSPLDGG LLTIRGRNLGRRLSDVAHGWWIGGVACEPP DRYTVSEBIVCVTGPAPGLSGVVTVNASK GKSDRFSVVLPLVHSLEPTMGPKAGGTRI HONDLHVGSEQUVLNDTDPCTFLMKTDT ACTMPEGALPAPVPVCVREERRGCVHGM: WYMQNSMAAVHHGREPTLCKVLNSTLITCF GALSNASAPVDFFINGRAYADEVAVAEELL PEBAQRGSRFRLDVLRVPQFSTKARKWK HFGPFLTLVHVSTKGAGKEQDSLGLGSE RKKGGVSCDIQVSDRIHCSVRESUGAVV LPITIQVGNFNQTIATLQLGGSETAINISVIVC LLLLSVVALFVFCTKSRRAERFWGKTLLQB BMESQREEIRKGFAELGTDMTDLTREINR GIFFLEYKHFVTRIFFFRCSSVYEGRTVMSWILL LLSVVALFVFCTKSRRAERFWGKTLNWSKIC SGSSQAQETHPLLGEWKIPESCRPNME GISLPSSLDNKEIFFFRCSSVYEGRTVMWSKIC NSGGSSQAQETHPLLGEWKIPESCRPNME GISLPSSLDNKEIFFFRCSSVYEGRTVMWSKIC VSCLRETVGEPFFLLLCAIRQQINKGSIDATI KARYTLMEEWLLREHIBARFRRINVSFQG MDSLSVRAMDTDTLTQVKERLEAFCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKLNTLAHYKJPGGASLAMSLIDKK NTLGRYKDLDTEKHRSWFRENLTNVSFQG MDSLSVRAMDTDTLTQVKERLEAFCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKLNTLAHYKJPGGASLAMSLIDKK NTLGRYKDLDTEKHRSWFRENLTNVSFQG MDSLSVRAMDTDTLTQVKERLEAFCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKLNTLAHYKJPGGASLAMSLIDKK NTLGRYKDLDTEKHRSWFRENLTVSYFQG MDSLSVRAMDTDTLTQVKERLEAFCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKLNTLAHFUNSEQLEAFCRING PDTLHIWKTINSLPLRFWWILKNPQFVFDID THIDDGALSVIAAGPTAARTQLOKKDDJFTKYA RYRQIMAALEARRYQNFTHJNVSFQG MOSLSVRAMDTDTLTQVKERLEAFCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS MAHLLAEESRKYQNEYNKJQODMFTLSE MHAHLAEESRKYQNEYNKJQODMFTLSE MHAHLAEESRKYQNEYNKJQDLGKERGI PDTLHIWKTINSLPLRFWWILKNPQFVFDID THIDDGALSVIAADGLFRANKSTRLLOKDTSPL MAHL			l	1			
SDPPPGAQSYAYLALAINSEARAGDKESQAR. LARICLPHOAGGDAKKLTESYQIG.GCCAC AGRGDLYSRUSYPFARERLFAVPERPOAG ARAAPALCAFRADVRAAIRAARTACFV APDVVAVLDSVVGGTGPACERKLNIQLOP LDCGAAHLQHPISLQPLATPYPRAPGIT: AVASVNNYTAVFLGTVNGRLIKINLNESM VVSRRVVTVAVGEPVHHVMQFDPADSOY LATTSHQMARVKVAACNVHSTCGDCVGAA AYCOWCALETRCTLQQDCTNSSQQHFWTS SEGPSRCAMTVLPSEIDVRQPTFOMILQIS LPSLSGMEBAACDYGNNIRTVARVPGPAFGI AYCOWCALETRCTLQQDCTNSSQQHFWTS GRIVKANFTIYDCSRIAQVYPHTACTISCL QWPCFWSQQHSCCSNQSRCEASPNTTSK CPRTLSFLAVPTIGGSQNILVPLANTAFFG AALECSFGLEEIFEAWVNESVYRCDQVVI TTRKSQVFPLSLQLKGRPARFLDSPEPMTVV VNNCAMSSPDSCGLGREDLGHLCMWSD RLRGPLQPMAGTCPAPEIRAEPLSGPLDGG LLTIRGRNLGRRSDVAHGVWIGGVACEPL DRYTVSEEIVCVTOPAFGFLSGVVTVNASK GKSRDRSTVLPLHYBLSEPTMGPKAGGTRI HONDLHVGSELQVLNNDTDPCTELMRITDT ACTMPEGALPAPVQVVERFRRGCVHGNI: WYMQNPVTIAISPRSFNSGGRTTVAGGRR MYGNNSMAVHIGREFTLCKVLNSTLTCG GALSNASAPVDFFINGRAYADEVAVAEELL PEAQRGSRFRLDVLPNPQFSTAKREKWK HPGEPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGGVSCDQUSSDRIHCSVERSCHAWK LLLLSVVALFVECTSRRAERKYGKTLLOW LLLLSVVALFVECTSRRAERKYGKTLLOW GIPPLEVHHFVTRFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFRCSSLYERETYVLPS LNSGGSSQAGFHPLLGGWKPESCRPING GISPSSLDDNKEITPFRCSSLYERETYVLPS LNSGGSSQAGFHTLLGGWKPESCRPING GISPSLDDNKEITPFRCSSLYERETY ARTYTHEEWLARETTSPYCEM LNSGGSGAGRAG				<u>!</u>	sequence		
LARICLPHOAGGDAKKLTESYIQLGLQCAG AGRGDLYSRLVSYPARERIPAYPERPOGS ARAAPAALCAFRADVRAARAARTACFV APDVVAVLDSVVQGTOPACERKINQLOP LDCGAAHLQHEJSLQPLKATPVFRAPGIT AVASYNNYTAAVELTVINGEDVRAAGTAG VVSRRVVTVAVGEPVHHVMQFDPADSSYI LMTSHQMARVKVAACNVHSTCGDCVGAA AVCGWCALETRCTLQQDCTMSSQQHFWT SEGPSRCPAMTVLPSEDVRQETPGMILQIS LPSLSGMEMACQYGNNIRTVARVPOPAGIG IAYCNLLPEDQFPFFPRODNIVTVEMSVRV GRNIVKANFTITYDCSSTIAQVPHTACTSCL QWPCFWCSQQHSCVSNQSRCEASPPITTSPF CPRTLLSPLAPVTFIGSSQNILVPLANTAFFQ AALECSFGLEEFEAVWVNESVVRCDQVVT TTRKSQVFPLSLQLKGRPAFLDSPPMTVI VYNCAMGSPDCSQCLGREDLGHLCMWSDD RLROPLQPMAGTCPAPEIRAEIPLSGPLDGG LLTIRGRNLSDVAHGWWIGGVACEPP DRYTVSEEIVCVTOPAPGELSGVTVANSK GKSRDRFSYVLPLVHSLEFTMOPKAGGTRI HONDLHVOSELQVLVNDTDPCTFLMRTDT ACTMPEGALPAPVPVCVREERRGCVHGML. WYMQNYSMAVHHGREPTLCKVLNSTLITTC GALSNASAPVDFFINGRAVADESLI PEGAQRGSRFRLDVLPNPQPSTAKREWIK HPGEPTLTVHVSTKGAGKEQDSISLQSHE RVKIGQVSCDIQIVSDRIHCSVNSSLGSH RVKIGQVSCDIQIVSDRIHCSVNSSLGSH HPGEPTLTVHVSTKGAGKEQDSISLQSHE RVKIGQVSCDIQIVSDRIHCSVNSSLGSH GISHTSSLDNKHETPFKCSSLYERETVLPSS LINSQGSSQAGETHPLLGEWKIPESCRPINGE GISHTSSLDNKHETPFKCSSLYERETVLPSS LINSGGSSQAGETHPLLGEWKIPESCRPINGE GISHTSSLDNKHETPFKCSSLYERETVLPSS LINSGGSSQAGETHPLLGEWKIPESCRPINGE GISHTSSLDNKHETPFKCSSLYERETVLPSS LINSGGSSQAGETHPLLGEWKIPESCRPINGE GISHTSSLDNKHETPFKYCSSLYERETVLPSS LINSGGSSQAGETHPLLGEWKIPESCRPINGE GISHTSSLDNKHETPFKYCSSLYERETVLPSS LINSGGSSQAGETHPLLGARKRANDSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKINTLAHYKIPEGASLAMSLIDKK NTLGRVRDLDTEKTSVEKMLTNNSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKINTLAHYKIPEGASLAMSLIDKK NTLGRVRDLDTEKTSVEKMLTNNSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKINTLAHYKIPEGASLAMSLIDKK NTLGRVRDLDTEKTSVEKMLTNNSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKINTLAHYKIPEGASLAMSLIDKK NTLGRVRDLDTEKTSVEKMLTNNSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVDLEWFASSTGSYLIRBLDDTS VEDGRKKNVLDTERTSVEKMLTHRANGFROHDLOFTNNAMS GARTABATAGLACHTERTVEKMR RATTAGLOFT SATERTY MEMBERSTERTSVEKMENTRANSPGG MOSLSVRAMOTDTLTQVKEKILEAPCKNV SQWPRAEDVD				1			
AGRGLLYSRU-PARERLAYERFOOS ARAPALCARPRADVRAA(RAARTACFV APDVVAVLDSVVGGTOPACERKLNIOLOP) LDCGAAHLQHEISLQPLATFYPFRAGIT AVASV'NNYTAVEGPVHHVAQEDADSGY UNSRUVTAVGEPVHHVAQEDADSGY LMTSHQMARVKVAACN'HSTCGDCVGAA AYCGWCALETRCTLQQDCTNISSQGHFWT: SEGPSRCAMTVLYSEDWYRGPYGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAFGI AYCNLLPROQPPFPNQDHVTVEMSVRV GRNIVKANFTIYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCCNSQSRCEASPNPTSP CPRTILISPLAPVPTGGSQNILVPLANTAFFO AALECSFGLEEIFFAWWNESVKCDQVVI TTRKSQVFPLSLQLKGRPARFLDSPEPMTV VYNCAMSSPDCSCQCGREDGHCLMWSD RLRGPLOPMAGTCPAPERAGEPLSQPLDGG LLTRGRNLGRLSDVHGVWIGGVACEPL DRYTVSEEIVCVTGPAFGPLSGVVTVMASK GKSRDRFSVVLPLVHSLEPTMGPKAGGTRI HONDLHVGSELQVLNVDDTDCTELMRTDT ACTIMEGALPAVPVCVFERRGCVHONL WYMQNYSMAVHHGREPTLCKVLNSTLITCE GALSNASAPVDFFNGRAYADEVXAVEELL PEEAQRGSRFRLDVLPNPQPSTAKREKWIK HPGEPLTLVHIVSTKGAGKEQDSS,CLQSHE RVKIGGVSCDIQIVSDRIHICSVNESLGANV LPHTQVONFNQTIATLQLGGSETAINVSTVC LLLLSVVALFVFCTKSGRAERYWQKTLLQ) DEMESSICERENKGPGALQTDIMTDLTKELNR GIPPLEYKHFVTRTFPKCSSLYERTYVLSR GRISLSLTIALIBALEVSTSINKELLVDLI ASAARNPKLMLRYTSVVEKMLTNWMSIC YSCLRETVGEFFLLLCAKRQQINKGSIDATI KARYTTLREEWLLREHAGRERNAUNSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYDLDT VEDGRKKLNTLAHYKDFGASLAMSLIDKK NTLGRVRADDTEVHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLEQD KLLTVAREPETRKVORYYKQIQODMP1SE RCSLASLITALITAGTLQKFLDDI SYEDGRKKLNTLAHYKDFGGAEACRI PDTLHWKTNSLPLRFWVNILKNPQFVFDID TOHIDACSVILOAPGACSISDLQCKGEPC SQURAMCKGNQRNAGGSGPGALAPERNLELQD WANCKGNQRNAGGSGPGALAPERNLELQD WANCKGNGRNAGGSGPGALAPERNLELQD VANCKGNQRNAGGSGPGALAPERNLELD LKD VANCKGNGRNAGGSGPGALAPERNLELD LKD VANCKGNGRNAGGSGPGALAPERNLELD LKD VANCKGN				1			SDPPPGAQSYAYLALNSEARAGDKESQARSL
AGRGLLYSRU-PARERLAYERFOOS ARAPALCARPRADVRAA(RAARTACFV APDVVAVLDSVVGGTOPACERKLNIOLOP) LDCGAAHLQHEISLQPLATFYPFRAGIT AVASV'NNYTAVEGPVHHVAQEDADSGY UNSRUVTAVGEPVHHVAQEDADSGY LMTSHQMARVKVAACN'HSTCGDCVGAA AYCGWCALETRCTLQQDCTNISSQGHFWT: SEGPSRCAMTVLYSEDWYRGPYGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAFGI AYCNLLPROQPPFPNQDHVTVEMSVRV GRNIVKANFTIYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCCNSQSRCEASPNPTSP CPRTILISPLAPVPTGGSQNILVPLANTAFFO AALECSFGLEEIFFAWWNESVKCDQVVI TTRKSQVFPLSLQLKGRPARFLDSPEPMTV VYNCAMSSPDCSCQCGREDGHCLMWSD RLRGPLOPMAGTCPAPERAGEPLSQPLDGG LLTRGRNLGRLSDVHGVWIGGVACEPL DRYTVSEEIVCVTGPAFGPLSGVVTVMASK GKSRDRFSVVLPLVHSLEPTMGPKAGGTRI HONDLHVGSELQVLNVDDTDCTELMRTDT ACTIMEGALPAVPVCVFERRGCVHONL WYMQNYSMAVHHGREPTLCKVLNSTLITCE GALSNASAPVDFFNGRAYADEVXAVEELL PEEAQRGSRFRLDVLPNPQPSTAKREKWIK HPGEPLTLVHIVSTKGAGKEQDSS,CLQSHE RVKIGGVSCDIQIVSDRIHICSVNESLGANV LPHTQVONFNQTIATLQLGGSETAINVSTVC LLLLSVVALFVFCTKSGRAERYWQKTLLQ) DEMESSICERENKGPGALQTDIMTDLTKELNR GIPPLEYKHFVTRTFPKCSSLYERTYVLSR GRISLSLTIALIBALEVSTSINKELLVDLI ASAARNPKLMLRYTSVVEKMLTNWMSIC YSCLRETVGEFFLLLCAKRQQINKGSIDATI KARYTTLREEWLLREHAGRERNAUNSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYDLDT VEDGRKKLNTLAHYKDFGASLAMSLIDKK NTLGRVRADDTEVHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLYSPQG MDSLSVRAMDTDTLTQVKEKLEAPCKNW SQWPRAEDVDLEWHASSTGSYLLRENLEQD KLLTVAREPETRKVORYYKQIQODMP1SE RCSLASLITALITAGTLQKFLDDI SYEDGRKKLNTLAHYKDFGGAEACRI PDTLHWKTNSLPLRFWVNILKNPQFVFDID TOHIDACSVILOAPGACSISDLQCKGEPC SQURAMCKGNQRNAGGSGPGALAPERNLELQD WANCKGNQRNAGGSGPGALAPERNLELQD WANCKGNGRNAGGSGPGALAPERNLELQD VANCKGNQRNAGGSGPGALAPERNLELD LKD VANCKGNGRNAGGSGPGALAPERNLELD LKD VANCKGNGRNAGGSGPGALAPERNLELD LKD VANCKGN				1			LARICLPHGAGGDAKKLTESYIQLGLQCAGG
ARAPAALCAFREADVRAMARATIACTV APDVAVLDSVOGTOPACERKINIOLOP LDCGAAHLQHE.SILQPLKATPVFRAPGIT AVASVNYTA VFLGTVNGRLKINI.NESM VVSRRVVTVAYQEPVHHVMQEDPADSGYI LMTSHQMARVKVAACNVHSTCGDCVGAA AYCGWCALETRETLQQECTINSSQHFWT SEGPSRCPAMTVLSEIDVRQEYPGMILQIS LPSLSGMEMACDYGNNIRTVARVPOPAFGI LAYCNLLPRDQFPFFPPNQDHVTVEMSVRV GRNIVKANFTIYDCGSTNIQVYTHACTSCL QWPCCWCSQOHSCVSNQSRCEASPNTTSP CPRTLISPLAPVPTGGSQNILVYLANTAFFQ AALECSFGLEEFEAWWVRESVVRCDQVVI TTRKSQVFPLSLQLKGPARALDSFPPMTVI VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRCPLQPMAGTCTAPEERAEPLSGVLDGG LLTIRGRNLGRRLSDVAHOWMGGVACEPL DRYTVSEEIVCVTGPAFGHLSGVVTVNASK GKSRDRFSVVLPLVHSLEPTMGPKAGGTRI HONDLHVGSELQVLNDTDTPCTELMRTDTI ACTMPEGALPAPVPCVVRFERRSCVHGNL WYMQNFVTIAISPRRSPVSGGRTITVAGGRR MYONSMAVHINGEFTLCKVLNSTLITCE GALSNASAPVDFFINGRAYADEVAVAEELL PEAQRGSRFROUTVPPOFFSTAKREWMK HPGEPLTLVHVSTKGAGKEQDSLGLGSHE RVKJGQVSCDIQIVSDRIHCSVVRSEKGWKSTLLGAAV LPITIQVGNFNQTAATLGLGGSETAINVSIVIC LLLLSVVALFVFCTKSRRAREFWKGYTLLGU BMESSGREERKGFABLGTDMTDLTIKELNR GIFPLSYLHFVTRIFFFKCSSLYEERYYLFS LNSGGSSQAQETHLLGWKIPESCRPNME GISLFSSLLDNKHFILVGWKIFELGAAV LNSGGSSQAQETHLLGWKIPESCRPNME GISLFSSLLDNKHFILVGWKIFELGAAV KRYTLABEEWLLRENIEAKFRNLNYSFOG MOSLSVAMDTDTITLOKEKERLELVDLI ASAAKNPKLMLRRTESVVERMLTNWSMG CSLASLITIALHGRLLEYTSINKELLUNGSIDATI KARYTLABEEWLLRENIEAKFRNLNYSFOG MOSLSVAMDTDTITLOKEKERLELAFKK RQSHRKKLVTEINLAKFRRLNYSFOG MOSLSVAMDTTITLOKEKERLELAFKK RQSHRKKLVTEINLAKFRRLNYSFOG MOSLSVAMDTTITLOKEKERLELAFKK RQSHRKKLVTEINLAKFRNLNYSFOG MOSLSVAMDTTITLOKEKERLELAFKK RQSHRKKLVTEINLAKFRRLLOKSPOG MOSLSVAMDTTITLOKEKERLELAFKK RQSHRKKLVTEINLAKFRRLLOKSPOG MOSLSVAMDTTITLOKEKERLEAFKK RQSHRKKLVTEINLAKFRRLLOKSPOG MOSLSVAMDTTITLOKEKERLEAFKK RQSHRKKLVTEINLAKFRRLLOKSPOG MOSLSVAMDTTITLOKEKERLEAFKK RQSHRKKVTPEINLAKRPRLLOKSPOG MOSLSVAMDTTITLOKERLEAFCKN RQSHRKRVIPEINLAKRPRLLOKRPRF LLYNGGGARAR PTHLHWKTNSLPLRFRWANILKNPQFVFDII TOHIDACLSVIAQAFDACSISIOLJOCKOSP KLLYAKEPEVRKINGRYYKQIQOMTPLSG MYTHOMATATATATATATATATATATATATATATATATATATAT			1	l		ì	, ,
APDVAVLDSVVQGTGACREKLINIQLOP LDCGAAHLQHPISLQPILKAPRAPGIT AVASVNNYTAVPLGTVNGRLLKINLTREA VVSRRVYTAVVELGTVNGRLLKINLTREA VVSRRVYTAVGEPHHYMQFDPADSGYI LMTSHQMARVKVAACNVHSTCGDCVGAA AYCGWCALETRCTLQDCTNSSQQHFWT SGGPSRCPAMTVLPSEIDVRGYEPTGMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAG ILYCNLLPROPPEPPROMILQIS LPSLSGMEMACDYGNNIRTVARVPGPAG ILYCNLLPROPPEPPROMILQIS GRNIVKANFTIYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCSVNGSRCSAPPITSP CPRTLLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSQNILVPLANTAFFG CPRTLSPLAPVPTGGSGNILVPLANTAFFG CPRTLSPLAPVPTGGSGNILVPLANTAFFA CREEKERAWWSDCOCKGEEDLGHLCMWSD RLRGPLQPMAGTCPAPERAEPLSPPEDATV VVYNCAMGSPPCSCCLGEEDLGHLCMWSD RLRGPLQPMAGTCPAPERAEPLSPPEDATV VVYNCAMGSPPCSCCLGEEDLGHLCMWSD RLRGPLQPMAGTCPAPERAEPLSPPTSTAC GRSTPSTLYLPLYLPSLEPTAGETSCCCAPPT ACTIVE GRSTPSTLYLPLYLPSLEPTAGETSCCCAPPT ACTIVE GRSTPSTLYLPLYLPLSLEPTAGETSCCCAPPT GRASSASAPVPDFFINGRAYADEAGGTTIVAGER MYQNYSMAVHHIGREPTLCKVLNSTLITYC GRSTPAGTITATCGGSGTAINVSIVC LLLLSVVALFVGTKSRAERYWQNTLINGVL PPEDAGRGSRFRLDYLPNOFSTAKREKWK HPGGEPLTLVHYSTKGAGKCGSLGQSHE RVKIGQVSCDIQVSDRIHCSVVESLGASAV LPITIQVGSFNGTIATTQGGSTAINVSIVC LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWQNTLINGVL LLLLSVVALFVGTKSRRAERYWGNTLINGVSIC YSCLRETVGEPFFLLLCAIKQGINKSGDAIT KARYTINEBWLLRRNIEAFRNLNSFQG MSGLSVRAMDTDTLTQVKERLLAFCCSCL SCHRKKVLFBYLTRLLSTKGTLQKFLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLDDINS VEGGKKLNTLAHYKIPEGASLAMLD	_			1			
LDCGAAHLQHPLSILQPLKATPVERAPIGT AVASVNNYTVACHGTVNGRLIKNINEN WVSRRVVTVAYCEPVHHVMQFDPADSGY LMTSHQMARVKVAACNVHSTCGDCVGAA AYCGWCALETRCTLQQDCTNSQQLA AYCGWCALETRCTLQQDCTNSQQLA AYCGWCALETRCTLQQDCTNSQQLA SEGPSRCPAMTVLPSEIDVRQEPYGMILQS LPSLSGMEMACDYGNNRITVARVPGPAFG IAYCHLLPROQPPPPPROQPHVTVEMSVY GRNIVKANFTTYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCVSNQSRCEASPNPTSTS CPRTTLSPLAPPYTGSSQNILVPLANTAFFQ AALECSFGLEEIFEAVWVNESVVRCDQVVI TTRKSQVPFLS, QLKGRPARFILVPLANTAFFQ AALECSFGLEEIFEAVWVNESVVRCDQVVI TTRKSQVPFLS, QLKGRPARFILGSPLDGG ILTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEBIVCVTGFAFQPLSQVVTVNASK GESRPRSYVLPLVSHLEPTMGFKAGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPPVCVCRERGCVHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPPVCVCRERGCVHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPPVCVCRERGCVHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCVHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT GALSNAAPDVFTTMGFAGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCHGNL WYMQNPVITAISPRSSPVSGGRTITVAGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT GALSNAAPDVFTTMGFAGGTRI HGNDLHVGSELQVLVNDTDPCTELIMRTDT ACTMPEGALPAPVPCVCRERGCHGNL HTTTTTTTTT GALSNAAPVTTTTTTTTT CATTAGAT ACTMPTT ACTM				İ		ļ	
AVASVNNYTÄVPLGTYÜGRILLKINLINGEN VÖRRRVTVAVGEPYHYÜGRILLKINLINGEN VÄRRVTVAVGEPYHYÜGRILLKINLINGEN LAMTSHÖMÄRVKVAACNYHSTCODCÜGAA AYCGWCALETRCTILQQDCTNSSQQFWTS SEGPSRCPAMTVLPSEIDVRQEYPGMILQIS LEPSLSGMEMACDYGNNIRTVARVPOPAGE I LYSLSGMEMACDYGNNIRTVARVPOPAGE I LYSLSGMEMACDYGNNIRTVARVPOPAGE I LYSLSLAPPYTOGSONILVPLANTAFFG CPRILLSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGSONILVPLANTAFFG CPRILSPLAPPYTOGROPHEN TITKKSQVPFLSULGLGREARFLDSPEPMTVY VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGGLGPAWFUNGOVACED LLTIRGRNLGRRLSDVAHGVWIGGVACED LLTIRGRNLGRRLSDVAHGVWIGGVACED LLTIRGRNLGRRLSDVAHGVWIGGVACED LLTIRGRNLGRRLSDVAHGVWIGGVACED LLTIRGRNLGRRLSDVAHGVWIGGVACED MYONVSNAVHHIGREPTLCXVLNSTILTGE GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSRFRLDYLPNPOPSTAKREWIK HIGGEPLILVHYSTKGAGKEQDSLGLGSGT MYONVSNAVHHIGREPTLCXVLNSTILTGE GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSRFRLDYLPNPOPSTAKREWIK HIGGEPLILVHYSTKGAGKEQDSLGLGSGT MYONVSNAVHHIGREPTLCXVLNSTILTGE GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSRFRLDYLPNPOPSTAKREWIK HIGGEPLILVHYSTKGAGKEQDSLGLGSGT MYONVSNAVHHIGREPTLCXVLNSTILTGE GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSRFRLDYLPNPOPSTAKREWIK HIGGEPLILVHYSTKGAGKEQDSLGLGSGT MYONVSNAVHHIGREPTLCXVLNSTILTGE GALSNASAPVTTAGKER GYPTLEVHFFYTTFFKCSSLST,VERTYULSS LLSVALAPPTRIFFKCSSLLLQQKDFAVW RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASSAANNFALMLRRTESVVEKMITNWMSIG SISPSALDNAMLLAPPTARTAFTQLQKIDGNEN VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGRALNTLAHVAPPGGASLAMSLIDKS VEDGGRKLNTLAHVAPPGGASLAMSLIDKS VEDGGR						ł	
VVSRRVYTVAYGEPYHHVMQFDADSOCYGAA AYCGWCALETRCTLQQDCTNSSQQHFWTS SEGFSRCPAMTVLPSEIDVRQETYGMILQS LMTSHQMARW VAACAVHSTGGDCVGAA AYCGWCALETRCTLQQDCTNSSQQHFWTS SEGFSRCPAMTVLPSEIDVRQETYGMILQS LFSLSGMEMACDYGNNIRTVARVPGPAFGI IAYCHLLPRDQFPFPPPNQHYTVEMSVRV GRNIVKANFTIYDCSRTAQVYFTACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCL QWFCFWCSQQHSCVSNQSRCEASPNTTHACTSCC CPRTLLSPLAPPYTGSQSNLUPLANTAFFQ AALECSFGLEEIFEAVWVNESVYRCDQVYI TTRKSQVFPLSQLICREDHUMSD RLRGPLQPMAGTCPAFEIRAIEPLSGPLOGH ALECSFGLEEIFEAVWVNESVYRCDQVYI TTRKSQVFPLSQLICREMSSVAICHWWIGGVACEP, DRYTVSEBIVCVTGPAPGPLSGVTVTNASK GRSRFSSVLIPLVLPSLEPTTHGFAGAGTRI HGNDLHVGSELQVLVNDTDPCTELMRITDT ACTMPGGALPAPPVCVCREFGCVHGNI. WYMQNPVITAISPRRSPVSGGRTITVAGERI MVQNYSMAVHHIGREPTLCKVLNSTLITTC GALSNASAPVPSUPFINGRAYADEVAVAEELL PPEAQRGSRFRLDYLPNQFSTAKRRKW HPGGFLTLVHVSTRGGAKEQDSLGJUSHE RVKIGQVSCDIQVSDRIHCSVVESLGAGAW LPITIQVGNFNGTIATLQLGGSFTAINSIVL LLLLSVVALFVFCTKSRRAERYWQKTLLQW LPITIQVGNFNGTIATLQLGGSFTAINSIVL LLLLSVVALFVFCTKSRRAERYWQKTLLQW EMESQIREEIRKGFAELQTDMTDLTRELNR GISLFSSLLDNKHLIVFVHALEQQKDFAVR RCSLASLLTIALHGRLEYYTSIMKELLVYDL ASAAKNPKLMRRTESVVERMLTNWMSIG YSCLRETVGEPFFLLLCAIKQQNKGSIDATI KARYTLNEEWILLRENIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKERLAFCCNM SQWPRAEDVDLEWFASSTQSYNROLDDIS VEDGRKKLNTLERNIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKERLAFCCNM SQWPRAEDVDLEWFASSTQSYNROLDDIS VEDGRKKLNTLERNIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKERLEAFCCNM SQWPRAEDVDLEWFASSTQSYNROLDDIS VEDGRKKLNTLERNIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKERLEAFCCNM SQWPRAEDVDLEWFASSTQSYNROLDDIS VEDGRKKLNTLERNIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKERLEAFCCNM SQWPRAEDVDLEWFASSTQSYNROLDDIS VEDGRKLNTLATHATPEGASSALMSLIDK NTLGRVRDATFTTFFTCLATLATRATQUQHKFEQVW MEIDTYECYSEA POTHIWKTNSLFRYWMILARPROPVFDIE TDHIDACLSVIAQAFIDACSIDQLGKDSPF KLLYAKEBFSYRKLQRYRYGQQDMTPLS MAHALAEESRKTQNGFTNNVAMAEIYKYA RYRQMAALAEDATTARRQQLPROVSWEM MAHALAEESRKYQNGFPONSWEM MAHALAEESRKQNGFTONNAMAEIYKYA RYRQMAALAEDLAFRANTELQQ PSPLFHSHLETQLLCTRACPEQVW MEDDTYPCYSEA POST			1	!			•
LMTSHQMARVKYAACNYSTICGOUVGAA AYCGWCALETRCTLQQDCTMSSQQHFWT SEGPSRCPAMTVLPSEIDVRQEYPGMILQIS LIPSLSGMEMACDYGNNIRTVARVPGPAFGI IAYCNLLPRDGPPFPPDQHVTVEMSVRV GRNIVKANFTIYDCSRTAQVYPHACTISCL QWPCFWCSQQHSCVSNSQCASSPNTYSE CPRTILSPLAPVPTGGSQNILVPLANTAFFQ AALECSFGLEBIFFA WYNESVVRCDQVVI TTKKSQVFPLSLQLKGRPARFLDSPEPMTV VYNCAMGSPDCSQCLGREASPNTYSE CRTTLLSPLAPVPTGGSQNILVPLANTAFFQ AALECSFGLEBIFFA WYNESVVRCDQVVI TTKKSQVFPLSLQLKGRPARFLDSPEPMTV VYNCAMGSPDCSQCLGREGHLCMWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRTGSDVAHGVVBIGGVACEPL DRYTVSEEIVCVTGPAFGPLSGVVTVNASK GKSRDKFSYVLPLVHSLEPTMGFKAGGTRI HGNDLHVGSELQVLLVNDTDPCTELMRTDT) ACTMPEGALPAPVPCVRFERRGCVHGNL WYMQNPVTLAISPRSPSVGRTTVAGGRE MVQNVSMAVHHIGREPTLCKVLNSTLTTCP GALSNASAPDPFINIRGAPVAVAEELI PEAQRGSRFRLDVLPNPGPSTAKREKWIK HPGEPLTLVHIVSTKGAGKEQDSLGLOSHE RVKIGQVSCDIQIVSDRIHEIVSVHGAKAGKEQDSLGLOSHE RVKIGQVSCDIQIVSDRIHEIVSVHCLGGSETAINIVSVIC LLLLSVVALFVPCTKSRRAFWQKTLLQD EMESQIREEIRKGFAELQTDMTDLTKELNR GIPFLEYKHFVTRTFFPKSSLYEERYVLPSS LNSQGSSQAQETHPLLGGSKETAINIVSVIC LLLSVVALFYVTRTFFPKSSLYEERYVLPSS LNSQGSSQAQETHPLLGGSKETAINIVSVIC LLLSVALFYTSTRFFKSSLYEERYVLFSS GISFSSLLDNAHFILIVEVHTSKETRSLVFBLVDLI ASAAKPRLMALRATESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAKQNKGSIDATI KARYTLNEEWLLRENBEAKFRNLNVSPQG MDSLSVRAMDTDLTTQVKELLEAFCKNV SQWPRAEDVDLEWFASSTQSYLRDLDDTS VEGGRKLINTLAHYKNEGASLAMSLIDKK NTLGRVEDLDTEKYFPHLVPTDELEAFKS RQSHRKVLPETYTRILSTKGTLOKFLDDI KALLSREDKPPLAVKYFPDFLEDQAEKGG PPDTLHWKTNSLPLFRWVNILKNPQFVDLI TDHIBACLSVIAQAFIDACSISDLQLGKDSPP KLLYAKEPFSKRIVQRYYKQQDDWPLSE MNABLAGESRKYQNEFTNIVAAMAEIKYV RYRPQMAALEANPTARRTQLQHKFEQVV MEDNIYECVSEA MNABLAGESRKYQNEFTNIVAAMAEIKYV RYRPQMAALEANPTARRTQLQHKFEQVV MEDNIYECVSEA POSPLFHSHL ETLQLLRTAQLFEQVSWFWG VANCKGNQRNMGSPOPSLAFERNLELQN LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDTLRVHRSSTTSSLLD LGYSLLMGKLRRPVAKDT				1			
AYGGWCALETRCTLQQDCTMSSQQHFWTS SEGPSRCPAMTLPSEDVAGEYPGMILQIS LPSLSGMEMACDYGNIRTVARVPGPAFGI LYCNLLPRDGPPFPPPNQDHTVTEMSVRV GRNIVKANFTIYDCSRTAQYYPHTACTSCLL QWPCFWCSQQHSCVSNQSRCEASPNPTSPC CPRTLLSPLAPVPTGGSRJAVYPHTACTSCLL QWPCFWCSQQHSCVSNQSRCEASPNPTSPC CPRTLLSPLAPVPTGGSRJAVYPHTACTSCLL QWPCFWCSQQHSCVSNQSRCEASPNPTSPC CPRTLLSPLAPVPTGGSRJAVYPHACTSCLL QWPCFWCSQQHSCVSNQSRCEASPNPTSPC CPRTLLSPLAPVPTGCSRJAVYBCAVCDQVVI TTRKSQVPPLSQLAFTLDSPEPMTVI VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLQPMAGTCFAPEIRAIEPLSGPLOGO LLTIRGRNLGRRLSDVAHGWVIGGVACEPL DRYTVSEEINCVTGPAFGLSGVTVTNASK GKSRDRFSSVLPLVHSLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDFTCELMRIDT ACTMPEGALAPPVPCVSTERGCVHGNL: WYMQNPVITAISPRSSPVSGGRTITVAGER MVQNVSMAVHHIGREPTLCKVLNSTLITCF GALSNASAPVDFFNGRAYADEVAVAEELL PEEAQRGSRFRLDVLPNPGFSTAKREWIK HPGEPLIT.VHVSTKGAGREQDSLGLQSHEE RVKIGQVSCDIQIVSDRIHICSVNESLGAAV LPITIQVGNFNQTIATLQSERGQDSLGLQSHEE RVKIGQVSCDIQIVSDRIHICSVNESLGAAV LPITIQVGNFNQTATLQSERGUSSLGLASVL LLLSVVALFVPCTKSRRAERYWQKTLLQN EMSSQLERENGFABLQDTDLTKELNR GIPFLEYKHFVTRIFFPKCSSLYEERYVLPS LLNSQGSSQAQETHPLLGEWKIPESCRFMEE GISLFSSLLDNKHFLIVPVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMELUVDL ASAAKNPKLMLRATESVEMALTHVMSFQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYTLRDLDDTS VEDGRKKLNLRATERSVEMLTNWSSGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYTLRDLDDTS VEDGRKKLNLRATERSVELLVDLI KARYTLNEEWLJREENGRAKPNILNVSPGQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYTLRDLDDTS VEDGRKKLNLRATERSVLLEVINGRAKRENNLNVSPGQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEVTLARCHTSKGTLQKFLDDD KALLSIREDKPPLAVKYFPDFLEQAKKGIS PDTLHWKTNSLPLFREVNILLEPHACLEPHAC RGSHRKVLPEPTLTRILTKGTLQKFLDDD TDHIDACLSVIAQAFDACSISDLQLGKDSPPI KLLYAKEIPSPRKIVQRYYKQIQDMFLSE MNAHLALESSRKYQNEFTNIVAMAEIKYX RYRPQIMAALEANPTARRTQLQHEFGQVW MEDNIYECVSEA MNAHLAERSSKLLQVGYYKQIQDMFLSE MNAHLAERSSKLLQVGYYRDGKDGNSGSPRFSPLLAFFRILEI,OH LGYSLLMGKLRPVAKDTLRVHRDSTPSPL LKD 730 2080 A 5747 1 382 FIKKCMKKAFRSSKLLQVGYYPDGKDDVRW							
SEGPSRCPAMTULPSEDNRQEYPÖMILQIS LPSLSGMEMACDYGNHENTYARVPGPASCI IAYCHLLPRDGFPFPPNQDHVTVEMSVRV GRNIVKANFTIYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCVSNQSCASPNFYSPS CPKTLLSPLAPVPTGGSQNILVPLANTAFFQ AALECSFGLEEIFFAVWNESVVRCDQVVV TTRKSQVFPLSLQLKGRRARFLDSPEPPMTVV VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLQPMAGTCPAPEIRISPLSGPLDGG LLTIRGRNLGRRLSDVAHGWNIGQVACEPL DRYTVSEELVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVOSELQVL.VNDTCPELMRITDTI ACTMPEGALPAPVPVCVRFERRGCVHGNL, WYMQNPVITAISPRRSPVSGGRITTVAGER MVQNVSMAVHHIGREPTLCKVLNSTLITCF GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSFRLDVLPPNGSTAKREKWIK HPGEPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCOLQIVSDRIHEDCKVLNSTLITCF GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSFRLDVLPPNGSTAKREKWIK HPGEPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCOLQIVSDRIHEDCKVLNSTLITCF GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSFRLDVLPPNGSTAKREKWIK HPGEPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCOLQIVSDRIHEDCKVLNSTLITCF GALSNASAPVDFFINGRAYADEVAVAELL PEEAQRGSFRLDVLPPNGSTAKREKWIK HPGEPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCOLQIVSDRIHEDCKVINSELGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKSTLQSH GISLFSSLLDNKHFILIVFVHALEQQKDSCHSCH GISLFSSLLDNKHFILIVFVHALEQQKDFAVR GISLFSSLLDNKHFILIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNMSSIGG MDSLSVRAMDTDTLTQVKEKILEAFCKNV SQMPRAEDVDLEWFASTSVTRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVRLDLTEKKYFHLVLPTDELAPKKS RQSHRKK VLPEITYLTRLISTKGTLQKFLDD KALSIREDKPPLAVKYFFDFLEEQAKRGR PDDTLHIWKTNSLPLFRFWINKPOFVTPDIE TDHIDACLSVIAQAFIDACSISDLQGKDSPI KLLYAKEPEVRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKVQNEFNTNVAMAEIYKYA RYRPQMAALEANFTARRTQLQHKFEQVV MEDNTYECTYSRA 4 292 QPSPLFHSHLETIQLLTAQLFEQVSWPWG VANGKGNGNMGSPOPSLLAFERNLELQN LGYSLLMGKLRPVAKDTLRVHRDSTPSPL LKD LGYSLLMGKLRPVAKDTLRVHRDSTPSPL LKD LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD LGYSLLMGKLRPVAKDTLRVHRDSTPSPL LKD LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD						l	LMTSHQMARVKVAACNVHSTCGDCVGAAD
LPSLSGMEMACDYGNNIRTVARVPQAFGG IAYCNLIPROQPPFPPRODDHYTVEMSVRY GRNIVKANFTTYDCSRTAQVYPHTACTSCL QWPCFWCSQQHSCVSNQSRCEASPNYTSK CRPTILSPLAPVPTGGSQNTYLANTAFTO AALECSFGLEIFFAAWVNESVVRCDQVVI TTRKSQVPFLSLQLKGRRAPRIDSPEPMTYV VYNCAMGSPDCSQCLGREDLGHLGWMSD RLRGPLQPMAGTCPAPERIFLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAFGLSGVVTVNASK GKSRDRFSSVLPLVHSLEPTMGFKAGGTRI HGNDLHVGSELQVLVNDTDFCTELMRIDT ACTMPEGALAPPVPVCVRFERRGCVHGNL. WYMQNPVITAISPRRSPVSGGRTITVAGER MYQNVSMA WHHIGREPTLAWTSTGAGEQDSLGLQSNE RVKIGQVSCDIQIVSDRIHICSVNESLGAV LPITTQVGNFNQTATLQGSETAINVSIVC LLLLSVVALFVFCTKSRRAERYWGKTLLQN EMESQIREERIFKGFAELQTDMTDLTKELNR GIPFLEYKHFVTRTFPKCSSLYERRYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPME GISLFSSLDNKHFLIVPVHALEQQKDFAVR RCSLASLLTIALHIGKLEYYTSIMKELIVDLI ASAAKNPKLMLRRTESVVEKMLTNWSNIGC WSCLRETVGEPFFLLLCAIKQQINKGSIDATI KARYTLNEERULREDIRAGRNLNVSFGGG MDSLSVRAMDTDTLTQVKEKILEAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKUPFGASLAMSLIDKK NITLGRVKDLDTEKYPHLVLPTDELAPFKKS RQSHRKKVLPETYLTRLISTKGTLQKFLDDD KALSIREDKPPLAVKYFFDFLEQASKRGI PDTLHWKTNSLPLFREWINKPSPCFDE TDHIDACLSVIAQAFDACSISDLQLGKDSPP KLLYAKEIPFYRKIVQRYYKQIQDMTPLSE MNAHLAEESKKYQNLEFNITVAAMAEIYKYX RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECVSEA 4 292 QPSPLFHSHLETIQLLETAQLFEQVSWFWG WNANGKGNGNNGSSPOSLLAFERNLELQD LGYSLLMGKLRPVAKOTLRVHRDSTPSPL LKD 730 2080 A 5747 1 382 FIKKCMKKAFRSSKLLQVGYTPDGKDDYRW				!		}	AYCGWCALETRCTLQQDCTNSSQQHFWTSA
IAYCNLLPROGPPPPPNODHVTVEMSVRV GRNIVKANFTIYDCSRTAQVPHTACTSCL QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QPFRLSPLAPVPTGGSQNILVPLANTAFFQ AALECSFGLEEIFEAVWVNESVVRCDQVVI TTRKSQVPPLSLQLKGRPARFLDSPEPMTVV VYNCAMGSPDCSQCLGEDLGHLCMWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLPLVHSLEPTMGFKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRTDT ACTMPEGALPAPVPVCVRFERRGCVGNGIL. WYMQNPVTAISPRSPSVGGRTTVAGGR MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNQFSTAKREKWK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSDGTHPLSCMAWAEELLAGAW LPITLQVSGTRATALDSVATEWWITTSPKCSSLVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSDGTHPLSCMAWAEELLAGAW LLSVALFVFCTKSRRAERYWKILLVDL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQOINKGSIDATI KARYTLLAEWLLRENHEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKEKLLAFCKNVV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKDEGASLAMSLDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKX RYKTGVAYKQUGDMTPLSE MAHLAEESRKYQNEFNTNVAMAEIYKYA RYKTQMAALEANFTARRTQLQIKKFEQVV MEDNIVECYSEA QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFENLELQN VANGKGNQNMGSPGSPSLLAFENLELQN VANGKGNQNGNMGSPGSPSLLAFEN					i		SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS
IAYCNLLPROGPPPPPNODHVTVEMSVRV GRNIVKANFTIYDCSRTAQVPHTACTSCL QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QWPCFWCSQQHSCVSNQSRCBASPNTTSNC QPFRLSPLAPVPTGGSQNILVPLANTAFFQ AALECSFGLEEIFEAVWVNESVVRCDQVVI TTRKSQVPPLSLQLKGRPARFLDSPEPMTVV VYNCAMGSPDCSQCLGEDLGHLCMWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLPLVHSLEPTMGFKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRTDT ACTMPEGALPAPVPVCVRFERRGCVGNGIL. WYMQNPVTAISPRSPSVGGRTTVAGGR MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNQFSTAKREKWK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSHE RVKIGQVSCDIQVSBGTTLCSVVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSDGTHPLSCMAWAEELLAGAW LPITLQVSGTRATALDSVATEWWITTSPKCSSLVATEWWIK HPGFPLTLVHVSTKGAGKEQDSLGQSDGTHPLSCMAWAEELLAGAW LLSVALFVFCTKSRRAERYWKILLVDL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQOINKGSIDATI KARYTLLAEWLLRENHEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKEKLLAFCKNVV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKDEGASLAMSLDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKX RYKTGVAYKQUGDMTPLSE MAHLAEESRKYQNEFNTNVAMAEIYKYA RYKTQMAALEANFTARRTQLQIKKFEQVV MEDNIVECYSEA QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFERNLELQN VANGKGNQNMGSPGSPSLLAFENLELQN VANGKGNQNMGSPGSPSLLAFENLELQN VANGKGNQNGNMGSPGSPSLLAFEN							LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ
GRNIVKANFITYDCSRTAQVYPHTACTSCL. QWPCFWCSQQHSCVSQSRCEASPHPTSPC CPRTLLSPLAPYPTGGSQNILVPLANTAFFQ AALECSFGLEEIFEAVWVNESVYRCDQVVI TTRKSQVPFLSLQLKGPRAFFLDSFEPMTVI VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLQPMAGTCPAFRLASPERIAPIS.SGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLP.VHSLEPTMGPKAGGTRI HGNDLHVGSELQVL.VNDTDPCTELMRTDTI ACTMPEGAL PAPVPVCVRFERRGCVHGNI: WYMQNPVITAISPRRSPVSGGRTITVAGERI MVQNVSMAVHHIGREPTLCKVLNSTLITCC GALSNASAPVDFFRNGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTJLVHVSTSTGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIHCSVNESLGAAV LPITIQVOFNPAGTIATLQLGGSETAINSVIVIC LLLLSVVALFVFCTKSRRAERTWQKTLLQN EMESQIREEIRRGFAELQTDMTDLTKELNR GIFFLSYKHFVTRTFFFYCSSLVEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPIME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLL ASAAKNPKLMLRRTESVVEKMITNWMSIC YSCLRETVGEPFFLLLCAIKQONIKGSIDAT KARYTLNEEWLLRENIEAKFRNILNVSFQG MDSLSVRAMDIDTLTQVKEKILEAFCKNIV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHU-VFTDELAEPKKS RQSHRKKVLPEIVLTRELLSTKGTLQKFLDDL KAILSIREDKPPLAVKYFFDFLEEQAEKRGIF DTLHIWKTNSLPLREWVNILKNPQFVFDIL TDHIDACLSVIAQAFTDACSISDLQGKDSPY KLLYAKEBPFSKIVQNYKQQDMTPLSE MNAHLAEESRYQNGFNTINVAMAEIYKYA RYRPQIMAALEANPTARTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSFLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNORMGSPOPSLLAFERNLELQIN CHYPPERVENCYPORYVKQQDMTPLSE MNAHLAEESRYQNGFNTINVAMAEIYKYA RYRPQIMAALEANPTARTQLQHKFEQVV MEDNIYECYSEA VANGKGNORMGSPOPSLLAFERNLELQIN CHYPPERVENCYPORYVKQQDMTPLSE MNAHLAEESRYQNGFNTINVAMAEIYKYA RYRPQIMAALEANPTARTQLQHKFEQVV MEDNIYECYSEA VANGKGNORMGSPOPSLLAFERNLELQIN CHYPPERVENCYPORYVKQLORMSPSPL LKYLLENDAMAEIYKYA RYRPQIMAALEANPTARTQLQHKFEQVV MEDNIYECYSEA VANGKGNORMGSPOPSLLAFERNLELQIN CHYPPERVENCYPORYBLIFERNLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLELQIN CHYPPERVENCYPORYBLIFTINGLETHT			l				
OWPCFWCSQQNSCVSNQSRCEASPNTTSP CPRTILSPLAPVPTGSQNILVPLANTAFFQ AALECSFGLEEIFEAVWNESVVRCDQVVI TTRKSQVFPLSLQLKGRPAFFLGSPENTVVI VYNCAMGSPDCSQCLGREDLGHICAWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLPLVHSLEPTMGFKAGGTRI HGNDLHVGSELQVLVANDTDPCTELMRIDT ACTMPEGALPAPVPVCVRFERRGCVHGNL: WYMQNPVITAISPRRSPVSGGRTIVAGERI MVQNVSMAVHHIGGEPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPOPGSTAKREKWK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCIDIQIVSDRIHICSVNCSLGAAV LPITIQVONFNQTIATLQLGGSETAINVSTVC LLLLSVVALFVFCTKSRRAERYWQKTLLQD EMESQUREEIRGGFALDDMTDLTKELNR GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS LNSQGSSQAGBTHPLLGEWKIPSCRPMWE GISLFSSLLDNKHFLIVFVHALEQGKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLARGQNIKGSIDAIT KARYTLNEEWLLRENIEAKPRAINVSFQG MDSLSVRAMDTDTLTJQVEKHLEAFCKNV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHJULPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARSTLAHVKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHJULPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARLSTRAMDTDTLTQVKKHLEAPCKNV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHJULPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARLSTRAMDTDTLTQVKKHLEAPCKNV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHJULPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARLSTRAMDTDTLTQVKKHLEAPCKNV SQWPRAEDVDLEWFASSTQSVILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHJUPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARLSTRAMDTDTLTQVKKHLEAPCKNV SQWPRAEDVDLEWFASSTQSVILROLDTS VEDGRKKLNTLAHYKIPEGASLAMSLICKK RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KARLSTRAMDTDTLTQVKKHLEAPCKNV SQWPRAEDVDLEWFASSTQSVILROLDTS VEDGRKGRAMGTGSPQFQSLLAFERNILELQIN KARLSTRAMDTATATATQLQHKFGVVX MEDNIYECYSEA PSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGROMGSPGSPGSPQPSLLAFERNILELQIN LKD 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQUSWPWG VANGKGROMGSPGSPGSPGSLLAFERNILELQIN LKD VANGKGROMGSPGSPGSPGSLLAFERNILELQIN LKD VANGKGROMGNGSPGSPGSPGSLLAFENILELQIN LKD VANGKGR			İ				
CPRTLLSPLAPYPTGGSQNILVPLANTAFFO AALECSFGLEEIFEA WVNESVVRCDQVVI TTRKSQVFPLSLQLKGRPAFLDSPEPMTVI VYNCAMGSPDCSQCLGREDLGHLCMVSD) RLRGPLQPMAGTCFAPEIRAIEPLSGPLDGG LLTIRGRNIGRRLSDVAHGWNIGGVACEPL DRYTVSEEIVCVTGPAPEISAUFPLSQVLTVANG GKSRDRFSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRITDTI ACTMPEGALPAPVPVCVRFERRGCVHGNIL WYMQNPVITAISPRSSPVSGGRITVAGGRI MYQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEBAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTJUHVTSTRAGKEQDSLGLQSHE RVKIGQVSCDIQUVSDRIIHCSVNESLGAAV LPHTQVGOFFNOTIATLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELINR GIPFLEFYKHFVTRTFFFYCSSLVEERYVLPS LNSGGSQAQSTHPLLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELINR GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDINKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFVIFALESCRPINGE GISLFSSLDNKHFLLIVFPVIFALESCRPINGE GISLFSSLDNKHFLLIVFPTGELAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLINTLAHYKIPEGASLAMSLIDKK NTIGRVKDAMDTDLTJOVEKBLEAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLINTLAHYKIPEGASLAMSLIDKK NTIGRVKDLDTEKYFPILVFPTDELEQABKRGI PDTLHIWKTNSLPLRFWVNILKNPQFVFDII TDHIDACLSVIAQAFIDACSISDLQGKDSPIP KLLYAKEPPSKRIVQRYMKQIQDMTPLSE MNAHLAEESRYQNGFNITNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSBA T30 2080 A 5744 3 292 QPSPLFHSHLETIQLLITTAQLPECVSWPWG VANGKGNORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNGKGRORMGSPOPSPILAFERILELQIN CYNCHATION TOTA				Ì	i	1	
AALECSFGLEEFFAVWVNESVRCDQVVI TTRKSQVFPLSLQLKGRPARFLDSFEPMTVI VYNCAMGSPDCSQCLGREDLGHLCMWSD. RLRGPLQPMAGTCPAFEIRAEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEBIVCVTGPAPGPLSGVVTVNASK GKSRDRFSYVLPLVHSLEFTMGPKAGGTRI HGNDLHVGSELQVL_VNDIDDPCTELMRITDT ACTMPEGALPAPVPVCVRFERRGCVHGNL. WYMQNPVITALSFRSVSGGRITIVAGGER MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTL.VHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAHVSIVIC LLLLSVVALFVFCTKSBRAERAFWQKTLLQD EMESQUREERKGFAELQTDMTDLTKELNS. GIPPLEYKHFVTRTFFFKCSSLYERRYVLPSS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASILITIAHGKLEYYTSIMKELLVDLL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDATI KARYTLNERWLIRENEAKPRNLNVSFGG MMSLSVRAMDTDITLYVKEKLLEAFCKNVI SQWPRAEDVDLEWFASSTIQSYILRDLDDTS YEDGRKLNTLAHTQVKEKLLEAFCKNVI SQWPRAEDVDLEWFASSTIQSYILRDLDDTS VEGRKLNTLAHTQVKEKLLEAFCKNVI SQWPRAEDVDLEWFASSTIQSYILRDLDDTS VEGRKLNTLAHTQVKEKLLEAFCKNVI SQWPRAEDVDLEWFASSTIQSYILRDLDDTS TOHIDACLSVIAQAFIDACSISDLQLGKDSP KLLYAKEPEYRKIVQRYYKQGUQDMTPLSE MNAHLAEESRKYQNEFMTNVAMAEIYKYW RYRQDMAALEANPTARRTIQLQHKFEGVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHISHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMSSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 1531 LKD 1541 LKD 1552 LKD 1554 LKD	1	a de la companya de l					
TTRKSQVPPLSLQLKGRPARFLDSPEMTVV VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEBIVCVTGAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEBIVCVTGAPEIRAIEPLSGPLDGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPI DRYTVSEBIVCVTGAPEIRAIEPLTVAGARG GKSRDRFSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDDFCTELMRTIDT ACTMPEGALPAPVPVCVRFBRRGCVHGNL: WYMQNPVITAISPRSPVSGGRTITVAGERI MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPOFSTAKREKWIK HHGEFLTUHIVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGNIFNQTIATLQLGGSSTAIIVSTVIC LLLLSVVALFVFCTKSRRAERFWQKYTLLQN EMESQREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLFSS LNSQGSSQAQETHPLJGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQNKGSIDATI KARYTLNEEWLLBEMARPRILNVSFQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTIGQLQKHE SQWPRAEDVDLEWFASSTIGTLQKFLDDI KALSIREDKPPLAVKYFFDFLEEQAEKRGI PDTLHIWKTNSLPLRFWVMILKNPGFVFDII TDHIDACLSVIAQAFIDACISIDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFINIVAMAEITYKY RYRPQIMAALEANPTARRTQLQHKFEGOVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHBLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSOPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW			l	ļ		ļ	
VYNCAMGSPDCSQCLGREDLGHLCMWSD RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGG LLTIRGRNLGRILSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVYTVNASK GKSRDRFSVVLPLVBLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRTDT ACTMPEGALPAPVPVCVRFERGCVHGNIT WYMQNPVITAISPRSPYSGGRITIVAGERI MYQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTLVHVSTKGAGKEQDS.GLQSHE RVKIGGVSCDQIVSDRIHCSVNESLGAAV LPITJQVGNFNQTIATLQLGGSSTAITVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQQREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS LNSQGSQAQEFIPLLGEWKIPESCRPIME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLL ASAAKNPKLMLRRTESVVEMLTNWMSIC YSCLRETVGEPFFLLCAIKQQNKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKEKLEAFCKNVI SQWPRAEDVDLEWFASTIGSYILRDLDDTS VEDGRKLNTLAHVKPGGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPETYLTRLISTKGTLQKFLDDI KALISIREDKPPLAVKYFFDFLEQAEKKGI KARYTLNEEWLARVTPFDFLEQAEKKGI KGSHRKKVLPETYLTRLISTKGTLQKFLDDI TDHDACLSVIAQAFIDACSISDLQLGKDSP KLLYAKEBPEYRKIVQRYYKQQIQDMTPLSE MNAHLAEESRKYQNEPKTNVAMAEIYKY RYRQDMAALEANPTARTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKONQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD T31 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			
RLRGPLOPMAGTCPAPEIRAIEPLSOPLOGG LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRRSYVLPLVHSLEPTMGPKAGGTRI HGNDL+VGSELQVLVNDTDPCTELMRTDTI ACTMPEGALPAPVPVCVRFERRGCVHGNL WYMQNPVTIAISPRRSPVSGGRTITVAGERI MVQNVSMAVHIGREPTLCKVLNSTLITCE GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPPNPFSTARKEWIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQVSDRIIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQD EMESQIREEIRRGFAELQTMTDTLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSK LNSQGSSQAQETHPLLGEWKIPESCRPME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASALTIALMKELLVYTISIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETTVGEPFFLLLCAIKQQINKGSIDATI KARYTLNEEWILRENIEAKPRNILNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPGASLAMSLIDKK NTLGRVKDLDTEKYPHLVJPTDELAEPKKS RQSHRKWLPSTLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPGFVPDIC TDHIDACLSVIQAPIDACSISDLQLGKDSPY KLLYAKEIPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRCYQNEFNTNVAMAEITYYA RYRPQIMAALEARNFTARRTQLQHKFEQVV. MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIL LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FIKKMRKAFRSSKLLQVGYTPDGKDDVRW				i			
LLTIRGRNLGRRLSDVAHGVWIGGVACEPL DRYTVSEEIVCVTGPAPGPLSGVVTVNASK GKSRDRSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRTDT ACTMPEGALPAPVPVCVRFERRGCVHGNL: WYMQNPVITAISPRRSPVSGGRTITVAGERI MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIDJSORHIHCSVYESLGAAW LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRRGAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSC LNSQGSSQAQETHPLLGEWKIPESCRPIME GISLFSSLLDNKHFLIVFVHALEQQKDFAVW GISLFSSLLDNKHFLIVFVHALEQQKDFAVW RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLCAIKQQNKGSIDATI KARYTLNEEWLLRENIEAKPRILNVSFQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYLRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYPHLVLPTDELAEPKKS RQSHRKKVLPSTVLTRLLSTKGTLQKFLDDI KAALSIREDKPPLAVKYFFDFLEEQAEKRGI PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPY KLLYAKEIPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQI LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDVRW				1			,
DRYTYSEEIVCYTGPAPGPLSGVYTVNASK GKSRDRFSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVOSSELQVLVNDTIDPCTELMRITDT ACTMPEGALPAPVPVCVRFERRGCVHGNL: WYMQNPVITAISPRRSPVSGGRTITVAGERI MVQNVSMAVHHGREPTLCKVLNSTLITCE GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFALDVLPNPQPSTAKREKWIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGNFNOTIATLQLGGSETAITVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRIFFPKCSSLYERRYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPIME GISLFSSLLDNKHFLLVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDATI KARYTLNEEWILLRENIBAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTOSYLRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKVLPEFVLTRILSTKGTIQKFLDDD KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHHWKTNSLPLRFWNILKNPQFYFDIL TDHIDACLSVIAQAFIDACSISDLQLGKDSPJ KLLYAKEPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV. MEDNIYECYSEA 730 Z080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIL LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 LOSI A 5747 1 382 FIKCMRKAFRSSKLLQVGYTPDGKDDYRW							RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT
GKSRDRFSYVLPLVHSLEPTMGPKAGGTRI HGNDLHVGSELQVLVNDTDPCTELMRTDT ACTMPEGALPAPVPVCVRFERRGCVHGNL: WYMQNPVITAISPRSPVSGGRTITVAGERI MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPUPFFINGRAYADEVAVAEELL PEEAQRGSRFRLDYLPNPQPSTAKREKWIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDQIVSDRIHICSVNESLGAAW LPITIQVGNFNQTIATLQLGGSETAIIVSIVIQ LLLLSVVALFVFCTKSRRAERYWQKTLLQD EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFFRCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLLVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIG YSCLRETVGEFFFLLLCAIKQQINKGSIDAIT KARYTLNEEWILLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWFRAEDVDLEWFASSTQSYURDLDDTS VEDGRKKLNTLAHVKIPEGASLAMSLIDKK NITLGRVKDLDTEXYFHLVLPTDELAEPKKS RQSHRKKVLPETYLTRLSTKGTLQKFLDDI KAALSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPRFWVNLKNPGFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFINIVAMAEITYSV MRYPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIA LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW						ſ	LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP
HGNDLHVGSELQVLVNDTDPCTELMRTIDT ACTMPEGALPAPVPVCVRFERRGCVHGNL WYMQNPVITAISPRRSPYSGGRTITVAGER MVQNVSMAVHHIGREPTLCKVLNSTLITCP GALSNASAPVDFFINGRAYADEVAVAEELL PEEAQRGSRFLDYLPNPOPSTAKREKWIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIHICSVNESLGAAW LPITIQVGSHPNQTIATT.QLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR; GIPFLEYKHFVTFFFKCSSLYEERYVLPSK LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIAHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDATI KARYTLNEEWLLRENIEAKPRNLNVSFQG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYLRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDII TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQUDMTPLSG MNAHLAEESRKYQNEFITNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 OPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				Į			DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE
ACTMPEGAL PAPVPVCVREERRGCVHGNL: WYMQNPVITAISPRRSPVSGGRTITVAGERI MYQNVSMAVHHIGREPITLCKVLNSTLITCE GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPQFSTAKREKWI HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKJGQVSCDIQIVSDRIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS: LNSQGSSQAQETHPLLIGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDIL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NITLGRVKDLDTEKYFHLJVFTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFFDFLEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQQDMTPLSE MNHHLAEESRKYQNEPTNIVAMAEIYKVA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA MNHHLAEESRKYQNEPTNIVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETILQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				ĺ			GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI
ACTMPEGAL PAPVPVCVREERRGCVHGNL: WYMQNPVITAISPRRSPVSGGRTITVAGERI MYQNVSMAVHHIGREPITLCKVLNSTLITCE GALSNASAPVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPQFSTAKREKWI HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKJGQVSCDIQIVSDRIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS: LNSQGSSQAQETHPLLIGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDIL ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NITLGRVKDLDTEKYFHLJVFTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFFDFLEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQQDMTPLSE MNHHLAEESRKYQNEPTNIVAMAEIYKVA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA MNHHLAEESRKYQNEPTNIVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETILQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							HGNDLHVGSELOVLVNDTDPCTELMRTDTSI
WYMQNPVITAISPRRSPVSGGRTITVAGERI MVQNVSMAVHHIGREPILCVLNSTLITICE GALSNASAVDFFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTLVHVSTISTKGAGKEQDSLGLQSHE RVKJGQVSCDIQIVSDRIJHICSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQ) EMESQIREEIRKGFAELQTDMTDLTKELNR GIPFLEVKHFVTRIFFPKCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPINME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLM.RRTESVVEKM.TNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLINEEWLLIRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYLRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFPFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIL TDHIDACLSVIAQAFIDACSISDLQLGKDSPJ KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLALESRRYQNEFNTINVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW	.			1			
MVQNVSMAVHHIGREPTLCKVLNSTLITCE GALSNASAPVDFFNGRAYADEVAVAEELL PEEAQRGSRFRLDYLPPPQFSTAKREK WIK HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQ0 EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRIFFFRCSSLYEERYVLPSY LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLFIDELAEPKKS RQSHRKVLPETYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPESYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							1
GALSNASAPVDFINGRAYADEVAVAEELI PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERY WQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLII ASAAKNPKLMLRTESVVEKMLTNWMSIC YSCLRETYGEPFFLLLCAIKQUNKGSIDATI KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYPHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAALSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQAYFVGLQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV/ MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			1
PEEAQRGSRFRLDYLPNPQFSTAKREKWIK HPGEPLTLVIHVSTKGAGKEQDSLGLOSHE RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGNFNQTIATT.QLGGSETAIIVSIVIC LLLLSVVALFVPCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPPLEYKHFVTRTFPFKCSSLYEERYVLPSI LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNFKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVAAMDTDTLTQVKEKILEAFCKNVV SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLFYDELAEPKKS RQSHRKKVLPETYLTRLLSTKGTLQKFLDDD KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYTRKIVQRYYKQQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDINTYEYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			
HPGEPLTLVIHVSTKGAGKEQDSLGLQSHE RVKIGQVSCOLIQIVSDRIHCSVNESLGAAV LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFFKCSSLYEERYVLPSI LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLII ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAALSIREDKPPLAVKYFFDFLEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFFVFDIE AGALSIREDKPPLAVKYFFDFLEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TTHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV/ MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LKD T31 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW	1			1			ł
RVKIGQVSCDIQIVSDRIIHCSVNESLGAAV LPITIQVGMFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYPHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDID TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				i I			
LPITIQVGNFNQTIATLQLGGSETAIIVSIVIC LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQTDMTDLTKELNR GIPFLEYKHFVTRTFFPKCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKKGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				ľ			, , ,
LLLLSVVALFVFCTKSRRAERYWQKTLLQN EMESQIREEIRKGFAELQITDMTDLTKELNR: GIPFLEYKHFVTRIFFFKCSSLYEERYVLPSG LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLII ASAAKNPKLMLRRTESVVEKMLTNWMSIG YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDITDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHILAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			
EMESQIREEIRKGFAELQTDMTDLTKELNR: GIPFLEYKHFVTRIFFPKCSSLYEERYVLPS LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDATI KARYTLNEEWLLRENIEAKPRNLNVSFQG MDDSLSVRAMDITDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLFTDELAEPFKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDID TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 I 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW]			LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV
GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSC LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIET TDHIDACLSVIAQAFIDACSISDLQLGKDSPT KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV/ MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNIELQU LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				ļ.			LLLLSVVALFVFCTKSRRAERYWQKTLLQME
LNSQGSSQAQETHPLLGEWKIPESCRPNME GISLFSSLLDNKHFLIVPVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPT KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTINVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV/ MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPIFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				ŧ			EMESQIREEIRKGFAELQTDMTDLTKELNRSQ
GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRTTESVVEKMLTNUMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPP KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT
GISLFSSLLDNKHFLIVFVHALEQQKDFAVR RCSLASLLTIALHGKLEYYTSIMKELLVDLI ASAAKNPKLMLRTTESVVEKMLTNUMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPP KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW					1		
RCSLASLLTIALHGKLEYYTSIMKELLVDLII ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNYSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			
ASAAKNPKLMLRRTESVVEKMLTNWMSIC YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGC MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIC TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
YSCLRETVGEPFFLLLCAIKQQINKGSIDAIT KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
KARYTLNEEWLLRENIEAKPRNLNVSFQGG MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW	1						
MDSLSVRAMDTDTLTQVKEKILEAFCKNVI SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
SQWPRAEDVDLEWFASSTQSYILRDLDDTS VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDID TDHIDACLSVIAQAFIDACSISDLQLGKDSPJ KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVVA MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				l			
VEDGRKKLNTLAHYKIPEGASLAMSLIDKK NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPT KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHILAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVVA MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW					l		
NTLGRVKDLDTEKYFHLVLPTDELAEPKKS RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHILAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVVA MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDI KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEC MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW			ļ	1			NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH
KAILSIREDKPPLAVKYFFDFLEEQAEKRGIS PDTLHIWKTNSLPLRFWVNILKNPQFVFDIE TDHIDACLSVIAQAFIDACSISDLQLGKDSPI KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF
PDTLHIWKTNSLPLRFWVNILKNPQFVFDICT TDHIDACLSVIAQAFIDACSISDLQLGKDSPT KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW				1			KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD
TDHIDACLSVIAQAFIDACSISDLQLGKDSPIKLLYAKEIPEYRKIVQRYYKQIQDMTPLSEGMNAHLAEESRKYQNEFNTNVAMAEIYKYARYRPQIMAALEANPTARRTQLQHKFEQVVMEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWGVANGKGNQRNMGSPQPSLLAFERNLELQINLGYSLLMGKLRPRVAKDTLRVHRDSTPSPLLKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEG MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW]	į	•	
MNAHLAEESRKYQNEFNTNVAMAEIYKYA RYRPQIMAALEANPTARRTQLQHKFEQVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW	[
RYRPQIMAALEANPTARRTQLQHKFEQVVV MEDNIYECYSEA 730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWG VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
VANGKGNQRNMGSPQPSLLAFERNLELQIN LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD 731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW	730	2080	Α	5744	3	292	
LGYSLLMGKLRPRVAKDTLRVHRDSTPSPL LKD							VANGKGNQRNMGSPQPSLLAFERNLELQIMG
	ſ			.			LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT
731 2081 A 5747 1 382 FLKCMRKAFRSSKLLQVGYTPDGKDDYRW							
	731	2081	Α	5747		382	
. I I I I CHAVIHANNIANAIMHIDHANCHANA	'-'	2001	А	3/4/	1 1	302	· · · · · · · · · · · · · · · · · · ·
T T T T T T T T T T T T T T T T T T T							FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion LSFSLRSSRVSGRHWKNFALVPLLREASARD RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS
732	2082	A	5753	198	3	GEK AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
733	2083	A	5754	2	2223	PGKLQAQLPCPSPVRFTSARIPPASRPQTKS AAGPPGLEAEGRAPESAGPGPGGDAAETPGL PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL SVANCLGAQTVQAPAEPAAGKAEQGETSGR EAPEAPAVGREDASAEDSCAEAGASGAADG ATAPKTEEEEEEETAEVGRGAEAEAGDLEQ LNRTSTSTKSAKSGSEASASASKDALQAMILS LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM DFSSMELDEALRKFQAHIRVQGEAQKVERLIE AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG ADIPRELVVGIYERIQQKELKSNEDHVTYVTK VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV NKLQKQAAHQREVFLFNDLLVILKLCPKKKS SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA QGDPQSKQGSPTAKREAALRERPAESTVEVSI HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH
734	2084	A	5788	8	362	GHRYSSGSRSLV SSVMGDLVGQGLEEQIVARDENSWLIDGGTP IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG NLVVCLMVYQKAAMRSAINILLASLAFADM LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY SFMGILNTLRHNALRHSYPEGICLSQASKLGL MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC WAPFTTYSLVATFSKHFYYQHNFFEISTWLL WLCYLKSALNPLIYYWRIKKFHDACLDMMP KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH GLARRLTTARRPPASSEQAQQELFNELKPAV DGANFIVNHMRDQNNYNEEKDSWNRVART VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP QPFPGDPYSYNVQDKRFI
738	2088	A	5881	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG LLDSSKLCDYENRFNTSKGGELPDRPAGVGV YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

SEQ ID NO: of nucl- cotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WCSQGADCITPGLYAMVGAAACLGGVTRMT VSLVVIMFELTGGLEYIVPLMAAMTSKWVA DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL AMDVMKPRRNDPLLTVLTQDSMTVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI LFN
139	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP DQALQELRKVARINGHKEAKNLTIEVLMSSV KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA VDFLGRATTALLLSFLGRRTIQAGSQAMAGL AILANMLVPQDLQTLRVVFAVLGKGCFGISL TCLTIYKAELFPTPVRMTADGILHTVGRLGA MMGPLILMSRQALPLLPPLLYGVISIASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM
741	2091	А	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV
742	2092	Α	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC
743	2093	Α	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEK WDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM
745	2095	A	5970	413	856	GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL NFPRRAQGPTWRSLEVIQSTEL

No. of collection No.	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
costic content conte			nou				
Sequence							
1914 ng to first amino acid residue of peptide peptide peptide of peptide peptide peptide sequence 747			1				
Peptide				i .	· •		
Peptide of peptide sequence			i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
746			Ì		peptide	-	/=possible nucleotide deletion, \=possible
RCARHGACORSCLASOPPYCOWHISSRGCUND	1		i		sequence		nucleotide insertion
RGSGGTDVDQAGNQESMEHGDQQQAGTOSS SGFGBSAYGVRRDLPPASARSSYPPLILASV	746	2096	Α	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
SGFGBSAYGYRRDLPSASRRSVPIPLLLASY							RCARHGACQRSCLASQDPYCGWHSSRGCVDI
AAAFALGASVSGILVSCACRAHRRISGIDA TORPRISLISTA RIHGGGPEPPPSKDGIDA VOTPOLYTTILPPPGGVPPPSKDGIDA VOTPOLYTTILPPPGGVPPPSKDGIDA VOTPOLYTTILPPPGGVPPPSKDGALPTPESTPE LPVKHLRAAGDPWEWNORNNAKEGFGRSR GGHAAGGPARVLVRPPPGCAQAVEVTTIL EELLRYLHGGPSPRKGAEPPARLTSRALPPEP APALLGGPSRPHEGASPILLDVPSGCASA PARPALSAPPLGGGGRKILPTSGALPPEP APALLGGPSRPHEGASPILLDVPSGCASA PARPALSAPPLGGGGKKILTJCAPPEGYRG RALKKVDVEKVOLSLXPPLVGPSSKQAVPNG GRPNP			l				RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
TFGLPRELSLRSLARLHGGGEPPPPSKDGD			İ				SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
VQTPQI,YTTFLPPFBCLACLPTPESTPE			ŀ				AAAFALGASVSGLLVSCACRRAHRRRGKDIE
LPVKHLRAAGDPWEWNONINNAKEGPRISR GGHAAGGPAPTU.VRPPPEGQAVEVTIL EELLRYLHGPOPPRKGAEPPAPLTSRALPPEP APALLGPPPECASPRU.VRPPPEGQAVEVTIL EELLRYLHGPOPPRKGAEPPAPLTSRALPPEP APALLGPPPECASPRLDLVPPEGRCASA PARPALSAPAPRLGVGGGRI,PFSGHRAPPAL LTRVPSGGFSRYSGGFGKHLLYLGRPEGYKG RALKRYDVEKPQLSLKPPL VSSRQAVPRIG GRIVIF GRIVIN CRIV	1		J				TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
GGIAAGGPAPRULVPPPPGCGGQAVETTL EELLYLHGPOPPRKGAEPPALTSRALPPEP APALLGGPSPPPHECASPLRLDVPPEGRCASA PARPALSAPAPILGVGGGRILFSGHRAPPAL LTRVPSGGPSRYSGGGRILLFULGYPEGYRGA A	i		İ	İ			
EELLRYLHGPOPPRKGAEPPAPLTSRALPPER	1						
APALLGGPSPRPHECASPIRLDVPPGRCASA PARPALSAPPRLGGGRCAPFSGHRAPPAL LTRVPSGGPSRYSGGGRLPFSGHRAPPAL LTRVPSGGPSRYSGGGRLPFSGHRAPPAL LTRVPSGGPSRYSGGGRLPFSGHRAPPAL LTRVPSGGPSRYSGGGRLPFSGHRAPPAL LTRVPSGGPSRYSGGGRLPFSGHRAPPAL LTRVPSGGPSRYSGGPGKHLLYLGRPGGYRG RALKRYDVERPQLSLKPPLVGPSSRQVPNG RALKRYDVERPQLSLKPPLVGPSSRQVPNG RALKRYDVERPQLSLKPPLVGPSSRQVPNG RALKLGQSDCGLTVTFRGHTGGYTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCWDPVQ RVLFSGSSDHSVIMWDIGGRGTALCQGD RVLFGGSVGCHAPTHOSK RVLFGGVPPYVPGVGCGSGRSSPIP MGFEFEVRVCDSCHEAITDEERAPTATHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS CVLLAFIGVAGVITYLISDSALFVETLOFLAV LTRAMLGVPQVRDRIBNARRRSTTAADS KDEEVKVAPRRSFIDDFIHWQWSSTSDHY RVSGDAFKTAYFLLKGAPLQFSVCGLLQVLV LTRAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV LTRAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV LTRAMLGVPQLYRNHRHQSTSGMSIKLVS VKEAIIGYGKGEDYPHTGAPT RVFTAD RVFT	İ		ļ				
PARPALSAPAPRLGYGGRRLPFSGHRAPPAL LTRVPSGGPSRYSGGGPGKHLLYLGRPGGYRG RALKRYDVEKPQLSLKPPLVGPSSRQAVPNG GRNF 747 2097 A 5998 2 754 DHASLPCSWNHRFDVETRHVFIGDHSGQVTI LKLEGENCTLVTTFRGHTGGYTALCWDPVQ RVLFSGSSDSSVIMWDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVWMNMD VERQETPEWLDSDSCQKCOPPFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDEERAPTATHDSK HNIVHVHFDATRGWLLTSGTDCKVIKLWDMT PVVS	1						l
LTRYPSGGPSRYSGGPGKHLLYLGRPGYTRG RALKRYDVEKPQLSLKPPLVGPSSRQAVPNG GRFNF 2097	1		ł				•
RALKRYDVEKPQLSLKPPLVGPSSRQAVPNG GRFNF	1						
	1		l				
747]						1 .
LKLEGENCTI-VTTERGHTIGOVTALCWOPVO RVLFSGSSDHSVIMWDIGGRKGTALCWOPVO RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVWNMD VERQETPEWLDSDSCQKCDQFFFWNFKQMW DSKKIGLRQHHCRKCGKAVCCKCSSKRSIPL MGFEFEVRVCDSCHEAITDEERAPTATHDSK HNIVHVHFDATRGWLLTSGTDVIKLLWDMT PVVS 748 2098 A 6001 2 747 AMVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRFESPLLWQSAIMIL TMLLMLKLCTEVRVANELINARRRSFTAADS KDEFVKVAPRSFLDPHIFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTSAMLGVPQLYRNHRHGSTEGMSIKMVLM WTSCDAFKTAYFLLKGAPLQFSVCGLLQVLV DLALLQQAYAPARHPQKPAPHAVHPTGTKAL THANLGYPQLYRNHRHGSTEGMSIKMVLM WTSCDAFKTAYFLKGAPLQFSVCGLLQVLV DLALLQQAYAPARHPQKPAPHAVHPTGTKAL GRPPDSELVRMHILEETFAEPSLQATQMKLK RARLADDLNEKLAGRPGPMELVEKNILPVDSS VKEALIGVGKEDYPHTQGFSSFDEDSDALDSS VKEALIGVGKEDYPHTQGFSFEDESDALDS DQPASQESQGSAASPSERVSESPSPVTINTP AQFASVSPTVPEFLKTPPTIAD QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRAVAKKLEKLLKRTQRAIABLIRERLK GQEDSLDSAVDAATEHKTC TDQAKVDNOPEELVRSAEDUSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGSLEVKGNIQFAIEYYEIKEL HVFVAQCKDLAAADVKKQRSDPVKAYLLE DKGKMGKKKTLVVKKTLNPVYNEILRYLSKEL HVFVAQCKDLAAADVKKQRSDPVKAYLLE DKGKMGKKKTLVVKKTLNPVYNEILRYLSKEL HVFVAQCKDLAAADVKKQRSDPVKAYLLE NRGEMKLALQYVPEPVPGKKLPTTGEVHIUV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKKTLTVVKYLRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIUV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKKTLTVVKYDGFREDLMAC VELTVWDHYKLTNQFLGGLRIGGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK KEIFSPFELLSVKPLLLLGVTCSQSMAFEELL SQUGGLGRFQMLHLVYHJPSLMLLIPHILLEFNF	747	2007		5000	2	754	
RVLFSGSSDHSVIMVDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVVWNMD VERQETPEWLDSDSCQKCDQFFFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSKRSSIPL MGFEFEVRVCDSCHAITDERAPTATFHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS 748 2098 A 6001 2 747 ANVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRFFSSPLLWQSAIMIL TILLMLKLCTEVRVANELNARRSFTAADS KDEEVKVAPRRSFLDFDPHHFWQWSSFSDVV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYKNHRRIQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHVPGTKALD LTEAMLGVPQLYKNHRRIQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHVPGTKAPL TAMLGVPQLYKNHRRIQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHVPGKAPHAVHPTGTKAL TJEAMLGVPQLYKNHRRIQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHVPGKAPHAVHPTGTKAL TJEAMLGVPQLYKNHRRIQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHVPGKAPHAVHPTGTKAL TAGMSVSTVPEFLKYPTAD TJEAMLGVFQLYSVCGLLQKVFGSPSPVTTNTP AQFASVSTVYPEFLKYPTAD TJEAMLGVFQLYSKGSFSPVTTNTP AQFASVSTVYPEFLKTPTAD TJEAMLGVFQLYSKGSFSPVTTNTP AQFASVSTVYPEFLKTPTAD TJEAMLGVFQLYSKGSFSPVTTNTP AQFASVSTVYPEFLKTPTAD TJEAMLGVFQLYSKGSFSPVTTNTP WDLKRDVAKKLEKLLKRTQRAIABLIRERLK GQEDSLOSAVDAATHKSTC TJEAMLGVFQLYSKGRSDPVTKAYLLFE TJEAMLGVFQLYSKGRSDPVTCTLLGVTCSQSMAFEELL TSCAMLGVFQLYSKGRSDPVTKAYLTE TJEAMLGVFQLYSKGRSDPVTCTLLFTLTCTCT TJEAMLGVFQLYSKGRSDPVTCTLTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	'*'	207/	^	שצבר		134	
DRVQALSYAQHTRQLISCGGDGGIVVMMD VERQETPEWLDSDSCQKCDQPFFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDEERAPTATFHDSK HNIVHYHFDATRGWLLTSGTDKVIKLWDMT PVVS 748 2098 A 6001 2 747 AMVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRRFESPLLWQSAIMIL TMLLMLKLCTEVRVANELNARRSFTAADS KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV QCVLAFTGVAGYITVLSIDSALFVETLGFLAV LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM WYSGDAFKTAYPLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL 749 2099 A 6002 2 447 GRPDRSELVRMHILEETFAEPSLQATQMKLK RARLADDLMEKLAQRFGPMELVEKNILFVDSS VKEAHIGVGKEDYPHTQGDFSFDEDSDALSP DQPASQESQGSASAPSPERVSESPSPVTINTP AQFASVSPTVPEFLKTPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRRQWFTPGFRLQAGGPCQEG WWELRINTYVPDEDLINKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRRPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNOPEKLVRSAEDVSTVPTQPDNPF SHPKLKRMSKSPALQESDDRETDTASE SYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDPGNLEVKGNIGPYGELYELLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEALLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVEAVLLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKKQRSDPYVAALLE HVFVAQCKDLAAADVKCRSDPYVAALLE HVFVAQCKDLAACH HVFVAQCKDLAACH HVFVAQCKDLAACH HVFVACACH HVFVACACH HVFVACACH HVFV							
VERQETPEWLDSDSCQKCDQPFFWNFKQMW DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDEERAPTATHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS 748 2098 A 6001 2 747 AMVFGGVVPYVPQYRDIRRTQNADGFSTYV CLVLLVANILRILFWFGRRFSPLLWGSAIMIL THLLMLKLCTEVPKORESPLLWGSAIMIL THLLMLKLCTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLWGSAIMIL THLMLKLTEVPKORESPLLVENGLOFSVORGLIV THLMLKLTEVPKORESPLLVENGLOFSVORGLIV THLMLKATEVPLUT THLMLKALTEVPLORE THE TOTAL THLMLKTEVPLORE THE TOTAL THLMLTEVPLORE THE TOTAL THLMLKTEVPLORE THE TOTAL THLMLTEVPLORE THE TOTAL THLMLTAN THE TOTAL THLMLKLTEVPLORE THE TOTAL THLMLTEVPLORE THE TOTAL THLMLTAN THE TOTAL THLMLKLTEVPLORE THE TOTAL THLMLTAN THE TOTAL THLMLKLTEVPLORE THE TOTAL THLMLTAN THE TOTAL THLMLTAN THE TOTAL THLMLTAN THE TOTAL THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTAN THLMLTHLLEPT THLMLTAN THLMLTAN THE TOTAL THLMLTAN THLMLTHLLTEN THLMLTHLLEPT THLMLTAN THLMLTHLLTEN THLMLTHLLTEN THLMLTAN THLMLTHLLTEN THLMLTAN THLMLTHLLTEN THLMLTAN THLMLTHLTHLLTEN THLMLTAN THLMLTHLTHLTHLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHLTHLTHLTHL THLMLTAN THLMLTHLTHTHLTHLTH THLMLTH				ļ			
DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL MGFEFEVRVCDSCHEAITDERAPTATHDSK HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS	1		ļ				
MGFEFEVR/CDSCHEAITDERAPTATTHDSK HNIVHVHFDATRGWILTSGTDKVIKLWDMT PVVS							
HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS							•
748 2098 A 6001 2 747 AMVFGGVVPYVPQYRDIRRTQNADGFSTYV	1					•	
CLVLLVANILRILFWFGRRFESPLLWQSAIMIL TMLLMKLCTEVRVANELNARRSSTAADS KDEEVKVAPRRSFLDFDPHFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYRNIRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLOFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL GRPDRSELVRMHILEETPAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTINTP AQFASVSPTVPEFLKTPFTAD DQPASQESQGSAASPSEPKVSESPSPVTTINTP AQFASVSPTVPEFLKTPFTAD ULTQAMLVLPHRPQWFTPGFRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKSPSSLTINLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVVGNIGFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKTLVVKKTLNPVYBEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNTMYYDGFRPEDLMEAC VELTVWDHYKLNOFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK TVLLIKGER TSST 2102 A 6028 108 1283 KEIFSPFELISVKPILCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHLLENF							
CLVILLVANILRILFWFGRRFESPLLWQSAIMIL TMLLMIKLCTEVRVANELNARRRSFTAADS KDEEVKVAPRRSFLDFPHHFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYRNIRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL 749 2099 A 6002 2 447 GRPDRSELVRMHILEETFAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDEDLINKRRVPQAKPDAV QEKVKEQLEAAKPPVIEVVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVVGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYBEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMYYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK TVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK SQVGGLGRFQMLHLVFILPSLMLLIPHLLENF	748	2098	Α	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL GRPDRSELVRMHILEFTFAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTT AQFASVSPTVPEFLKTPFTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERIK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNFF SHPDKLKKMSSSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIGFAIEYVESLKEL HVFVAQCKDLAAADVKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVFGKKLPTTGEVHWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELSUKPILCLLIGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF						1	
QCVLAFTGVAGYITYLSIDSALFVETLGFLAV LTEAMLGVPQLVRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL 749 2099 A 6002 2 447 GRPDRSELVRNHILEETFAEPSLQATQMKLK RARLADDLNEKIAQRYGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTYPEFLKTPPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSSYPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPICLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF] .						TMLLMLKLCTEVRVANELNARRRSFTAADS
TEAMLGVPQLYRNHRHQSTEGMSIKMVLM WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGKAL 749 2099 A 6002 2 447 GRPDRSELVRMHILEETFAEPSLQATQMKLK RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEMEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTNTTP AQFASVSPTVPEFLKTPPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGRRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TVDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	1						KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV DLAILGQAYAFARHPQKPAPHAVHPTGTKAL							QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
DLAILGQAYAPARHPQKPAPHAVHPTGTKAL	1						
Total	ļ l						
RARLADDLNEKIAQRPGPMELVEKNILPVDSS VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD 750 2100 A 6004 2 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRRPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF							DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	749	2099	Α	6002	2	447	
DQPASQESQGSAASPSEPKVSESPSPVTTNTP AQFASVSPTVPEFLKTPPTAD 750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF							
AQFASVSPTVPEFLKTPPTAD							
750 2100 A 6004 2 427 LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG WRWELRLRNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	1 1						
WRWELRLNYVPEDEDLNKRRVPQAKPDAV QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	750	2100		6004	2	427	7.02
QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	130	2100	A	0004	2	421	
WDLKRDVAKKLEKLLKRTQRAIAELIRERLK GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKA YLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	1						
GQEDSLDSAVDAATEHKTC 751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF							
751 2101 A 6007 33 1280 TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPJFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF][- 1
SHPDKLKRMSKSVPAFLQDESDDRETDTASE SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	751	2101	A	6007	33	1280	
SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	'		ا	200,		.200	
VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		ì					
HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		į					
DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		l			l		
QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPILRGSHLNSFYKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	[[[
TWDWDNKQNKQLRWYPLKRKTAPVALEAE NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF							
NRGEMKLALQYYPEPVPGKKLPTTGEVHIWV KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		1			ļ		•
KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		i		1	. !		NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	j 1		ļ		- 1		
EVDWMDSTSEEVALWEKMVNSPNTWIEATL PLRMLLIAKISK 752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF							
752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		ł					VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
752 2102 A 6028 108 1283 KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF		ŀ					
SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF	L1						
	752	2102	Α	6028	108	1283	
AAAIPGHRCWVHMLDNNTGSGNETGILSEDA		1			1	1	
	L						AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK
753	2103	A	6043	1	1470	KNLKEKA DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	Α	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRFTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE- GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106	A	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	Α	6063	54	419	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

NO: of NO: of No: of not n	CEOTE	Lecon	Mat	SEC	Dundicted	Deadleted and	L Amino anid company (Am Alamina C. Charlet
Decidide Seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
		1	nod				1
Sequence		1					
Bene							
mino acid recided sequence		uence	1				
Peptide Pept	uence		ļ	914	, -		
Peptide		ł	1	I	1		
		l	1		ž.	sequence	
	ĺ	1	1	1	peptide		/=possible nucleotide deletion, \=possible
1758					sequence		nucleotide insertion
PSDPSYEDMREVVCVKRIRPIVSNRWINSDEC LRAVILLIMSECWAHNPASRLTALRIKKTLAK MYESQDVKI 759 2109 A 6072 3 650 PGRRFRPAALEERAMEKLREKVPFQNRGKGT LSSIIPNNSDTRKATETTSLSSKPEYVNPDFFKW SKDPSSKSGNLETSEVGWTSNPEELDPIRLA LLGKSSGLSQVGSATSHPWSQCPPIEDQRISI KDKSTAGREFSGQVSHQTTSENQCTPPSSTV HSSVADMQNMPAAVHALTQPSLSAAPFAQ RYLGTLPSTGSTILPQCHAGNATVW FSQVADMQNMPAAVHALTQPSLSAAPFAQ RYLGTLPSTGSTILPQCHAGNATVW PRITTHMEPULLIGLRORIQLEACGMRRSSLLTRK VICKSDAPTGDVLLDEALKHVKETOPPETVQ NWIELLSGETWIPLELHYQLRONTGELAKHVKETOPPETVQ NWIELLSGETWIPLELHYQLRONTGELAKHVKETOPPETVQ NWIELLSGETWIPLELHYQLRONTGELAKHVKETOPPETVQ NWIELLSGETWIPLELHYQLRONTGELAKHVKETOPPETVQ NWIELLSGETWIPLELHYGNATVRQLL LDPEVECKANTNEVLWAWWAAFTK LLDPEVECKANTNEVLWAWWAAFTK PG 12111 A 6678 833 390 IVSFHLSGFRKFVKPFSFLSVHOLQVDEYHISV HQKLSADMADINSHLISLLUGAEDAGLARLMRD MKTTMKSRYMELYDLNGSLLNGKKRWNNH TELLGNIKANVQAJQRAGRLRVCKKPKNQVIT ACRDAIRSNINITLFKIMRVCTASS PSPLSSHALLSGARGAGRLRVGKKRRNQVIVRELEEAT ACRDAIRSNINITLFKIMRVCTASS PSPLSSHLINGSKAMALBGGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG SGGINASNINQLAKKVRLRKYEAKRRIANLKI QLAKLDSEAWPGVLDSEDBRILLINGKEELLK EMFFISPRKVTGGEVEQLEAARKRIANLKI QLAKLDSEAWPGVLDSEDBRILLINGKEELLK EMFFISPRKVTGGEVEQLEAARKRIANLKI QLAKLDSEAWPGVLDSEDBRILLINGKEELLK EMFISPRKVTGGEVEQLEAARKRIANLKI QLAKLDSEAWPGVLDSEDBRILLINGKEELLK EMFISPRKVTGGEVEQLEAARKRIANLKI QLAKLDSEAWPGVLDSEDSBAVGDSSGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS						-	QILTMLLRSLQQPSASWPRDCSSSCSW
RAVILKIMSECWAHNPASRITALRIKKTLAK MYESQOWKI	758	2108	Α	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
	}		1	İ	}	ļ	PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
		į .	1	ł			LRAVLKLMSECWAHNPASRLTALRIKKTLAK
759 2109 A 6072 3 650 FGRRFRPAALEERAMEKLEKVPFONGKUNG		1					
LSSIIPMNSDTRKATETTSLSSKEPVNDTPELDPIRLA	759	2109	Α	6072	3	650	<u> </u>
SKDPSSKSGNLLETSEQGWISNPEELDPIRLA LLGKSGLSCQVGSATSHPVSCQEPIDEDQRIST KDKSTAGREFSGQVSHOTTSENQCTPIPSSTV HSSVADMQNMPAAVHALLTQPSLSAAFFAQ RYLGTLPSTGSTTLPQCHGNATVW 760 2110 A 6077 3 730 PLRITIMEEVILLIGLKDREGYTSEWNDCISSC LRGGMLELPLRGRIQLEAGGMRKSSLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPIKLHVGYRERLAKIN VEKGVLITEKQNPILLFDMTTHPLTNNNIKQR LIKKVQEAVLDKWNDDPHRSRILALIYL AHASDVLENAFAPLLDEQYDLATKRVRQLLL LDPEVECLKAMINEVLMAVAFIK LIKKVQEAVLAKINEVLMAVAFIK LIKKVQEAVLAMINEVLMAVAFIK AHASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLDEQYDLATKRVRQLLT LAMASDVLENAFAPLLTDEQYDLATKRVRQLLT LAMASDVLENAFAPLLTDEQYDLATKRVRQLLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQYDLATKRVRQLT LAMASDVLENAFAPLTDEQX LAMASDVLENAFAPLTDEQX LAMASDVLENAFAPLTDEQX LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFAPLTDEAFAPLT LAMASDVLENAFATAR LAMASDVLENAFATAR LAMASDVLENAFATAR LAMASDVLENAFATAR L			1		i	***	1
LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISI KDKSTAGREFSQGVSHQTTSPSSTV HSVADMQNMPAAVHALLTQPSLSAAPFAQ RYLGTLPSTGSTITLPCHAGNATVW HSVADMQNMPAAVHALLTQPSLSAAPFAQ RYLGTLPSTGSTITLPCHAGNATVW WSVCSNDAPTGDVLDEALKHVKETQPPETVQ NWIELLSGETWIPLKH.HYQLRVRERERLAKIN VEKSQLTENGVLLEDALKHVKETQPPETVQ NWIELLSGETWIPLKH.HYQLRVRERERLAKIN VEKGVLTTEKQNLLLDTHTIPLTNINIKOR LIKKVQEAVLDKWVNDPHRMDRILLALIYL AHASDVLENAFAPLLDEGVALTKEVRQLLE LDPEVECLKANTNEVLWAVVAAFTK HQKLSADMADNSINLISLVATKEVRQLLE LDPEVECLKANTNEVLWAVVAAFTK HQKLSADMADNSINLISLVATKVARQLLE LOPEVECLKANTNEVLWAVVAAFTK HQKLSADMADNSINLISLVATKVARQLIV AKTOMASSYSTEM HQKLSADMADNSINLISLVAMIN TELLGINLKAVVQAQRAGRAGARALMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGINLKAVVQAQRAGRAGARALMRD KANTAMSRYMELYDLNRDLLNGYKIRWNNH TELLGINLKAVVQAQRAGRAGARAKIN HQKLSADMADNSINLISLVAMIN KORTASS HQKLSASSYSTEMSSPLPKQYLDVSSOTDISG SFGINSNNQLAEKVRLRRYEEAKRIANLKI QLAKLDSEA WPGVLDGSERDRILLINEKEELLK EMREISPRKWTQGEVEQLEMARKRLERLDLQ AARDTQKSSSSLESSSPLPKQYLDVSSOTDISG SFGINSNNQLAEKVRLRRLYEEAKRIANLKI QLAKLDSEA WPGVLDGSEDRILLINEKEELLK EMREISPRKWTGGEVEQLEMARKRLERLDLQ SKVEFLLLEGATGREPSGCTTIHIEDEVAKTQ KAEGGGRUARSLSGTFKSMTSLSPRSSLSS GSLVASSLDSSTSASFTDLYSPEQLISEL SKVEFLLEGATGREPSGCTTIHIEDEVAKTQ KAEGGGRUARSLSGTFKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCESLGNSAGERVRLEFETGKQLGQAV NTAQGCGLKVACVSAAVSDESVADDSOVYE ASVQRLARSLSGTFKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCESLGNSAGERVRLEFTGKQLGQAV NTAQGCGLKVACVSAAVSDESVADDSOVYE ASVQRLARSLSGTFKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCESLGNSAGERSTRWYNLLSYKYLKKQS RELKPVGVWAPASGPASTICLDSWRYPETISPRAVAE EEEEEVEEEGEGDVTFTEKASPDMOGYPALK VDEKNKGPALLQQADQKVNIR VAVLPCSESTTCLFRTPLDASDTLVTROFHEWST VSMSYPALLHQAMASGPASTRUARS PDMOGYPALK VDEKNKGPALHQANSKKRIPPPQPSSVK SLESSELIRTSJDLELDLQARTNYHSQLTGEIS VLKELFKPVTREPFW VSMSYPALHQAKSKKRPPPPQPSSVK SLESSELIRTSJDLELDLQARTNYHSQLTGTIS VLKELFKPVTREPFW MIPALSADDV PHPFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTGKMCLVPGKTKRLLFKFVAKTED VGKKIEITSVDLALGRETGRCVVLNWQGGGG DAASSCGEAPGAARSSKRENDLYDNEVHWOSIG DAASSCGEAPGAARSSKRENDLYDNEVHWOSIGGGG DAASSCGEAPGAARSSKRENDLYDNEVHWOSIGGGG DAAS	١.		ł	1			
CDKSTAGREFSGQVSHQTTSENQCTPPSSTV			Ì				
HISSVADMONMPAANHALLTOPSLSAAPFAQ RYLGTLPSTGSTILP,QCHAGNTVW 760 2110 A 6077 3 730 PLRLTLMEEVILLGLKDREGYTSFWNDCISSC LRGCMLIELPLRGRLQLEAGCMRRKSLTRK VICKSDAPTGDVLLDEALKHVEETQPPETVQ NWIELLSGETWNPLKLHVQLRNVEERLAKNI VEKGVLTTEKQNPLLLDAMTHPILTNINNIKQR LIKKVQEAVLDKWVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEGYDLDTAKRVRQLLLIY AHASDVLENAFAPLLDEGYDLDTAKRVRQLLALIYL AHASDVLENAFAPLLDEGYDLATKRVRQLLAL LDPEVECLKANTNEVLWAVVAAFTK 1 1 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLVQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGREVGKPKNQVTT ACRDAIRSNNINTLFKIMRVGTASS KKATICGEKEKQDLIKSLAMIKDOFFRTDRGS SFGRISNNQLAEKVARLIK VEPEAKRIANLKI QLAKLDSEAWPGVLDSERDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLENDLQ AARDTQSKALTERLKINSKRNQLVRELEEAT RQVATLHSQLKSLSSSSMOSLSSGSSPGSITSSR GSLVASSLDSSTSASTFDLYTYPPFEQLDSELQ SKVEFLLLEGATGRPSGCTTHHEDEVAKTQ KAEGGGRLQALRSLSGTFKRTSLSPRSLSS SFSPPCSPLMADPILLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERVRLEEPTGKQLGQAV NTAQGCGLKVACVSAAVSDESVAODSGVYE ASVQRLGASBAAAFDSDESEAVGATRIQIALK YDENKQFALIAQISNISTSKSLSSSSSSSSSSSSSSSSSSSSSSSVASLBSSTSSTTLEFRIFPLDASDTLVRNEVFW VSMSYPALHQAKTLRVDVCTTORSHLEECLGG AQISLAEVCRSGERSTRWYNILSYKYLKKQS RELKPYGVMAPASGPASTDASALLEQTAVE LEKRGGGRSTOTILEDSWRYFETISPHAVAE EEGEEVEEEGGEDOVTFEKASPDMOGYPALK VDKETNTETPASSPTVVRPCDRRVGTPSQGPF LRGSTITRKKTFSPOPQSQVVCRLNSSDSSTST LSKKPFFVRNSLERRSVRMKRSPPPPQSSVV SLRSSELIRTSJDLELDLQATRTWHSQCTGES VLKEKKPFFVRNSLERRSVRMKRSPPPPQSSVV SLRSSERLIRTSJDLELDLQATRTWHSQCTGES VLKEKKPFFVRNSLERRSVRMKRSPPPPQSSVV SLRSSERLIRTSJDLELDLQATRTWHSQCTGES VLKELKGRGRSTTOTLEGSKRKRSPFRPGSSPST LSKKPPFVRNSLERRSVRMKRSPPPPQSSVV SLRSSERLIRTSJDLELDLQATRTWHSQCTGES VLKELKGRGRSTGTLEGAKARAFTERP MMIPALSADDV PHRFSKLCVSFRNOEYNGFFYGFCYEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKEITSVDLALGRETGRCVVLNWQGGGG DAASSQEALAGARSSKREPLEDNEVHWSGGGG DAASSQEALAGARSSKRENPLDNEVHWOSH					i		
RRIGILPSTGSTILPQCHAGNATYW 760 2110 A 6077 3 730 PURITIMEEVILLIGLENGYTSFWNDCISSC LRGCMLIELPLRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVILDEALKHVKETQPETVQ NWIELLSGETWNPIKLIHPQLRNVRERLAKNI VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR LIKKVQEAVLDKWVNDFIMBORRILALIYL AHASDVLENARAPILDEGYDLATKRVRCILE LDPEVECLKANTINEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGABDARLMRD MKTMKSRYMELYDLNRDLINGYKRWNNH TELLGNLKAVNQAIQRAGRILRVGKFKNQVIT ACRDAIRSNINITLKIMKOTASS MKTMKSRYMELYDLNROKTASS MKTMKSRYMELYDLNROKTASS MKAITCGEKEKQDLIKSLAMLKDGFKTRGS HSDLWSSSSSLESSSFPLPKQVILDVSSQTDISG SFGNSNNQLAEKVRLRYEBAKRRIANKI QLAKLDSEAWPGVLDSEDDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKBLEKDLQ AARDTQSKALTERIKINSKRNQLVRELEEAT RQVATHASQLKSJSSSYGSGSSSSSSITSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGGFPSSELSGSSSPSGLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGGFPSSELSGSSPSPSLTSS GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGGFPSSELVGATRIQIALK YDEKNKQFALIJQLSNLSGSSPSPSLTSS SPSPCSPLMADPLLAGDAFLNSLEEPDFLSA TLCELSLGNSAGERYRLEEPGTEGKQLGQAV NTAQCCGLKVACVSAVSBESVAGATRIQIALK YDEKNKQFALIJQLSNLSTLYRLKYLEKQE SKVEFLLLEGATGGFPSEBAVGATRIQIALK YDEKNKQFALIJQLSNLSTLYRLKYLEKQE SKVEFLLEGATGGFPSEBAVGATRIQIALK YDEKNKQFALIJQLSNLSTLYRLKYLEKQE SKVEFLLEGATGGFPSEBAVGATRIQIALK YDEKNKQFALIJQLSNLSTLYRLKYLEKQE SKVEFLLEGATGFRSEBAVGATRIQIALK YDEKNKQFALIJQLSNLSTLYRLKYLEKQE LEKRQGGRSTTQTLEDSWRYEETSNEAVAE LEKRQGGRSTTQTLEDSWRYEETSNEAVAE LEKRQGGRSTTGTLEDSWRYEETSNEAVAE LEKRQGGRSTTGTLEDSWRYEETSNEAVAE LEKRQGGRSTTGVTLNSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSUDLQATRIVRSQLTTGSQGFP LRGSTIIESKTTSPGPQSVVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSUDLQATRIVRSQLTTPSQ TVIKELKEQLQAKSTGFREKMAFTTPR MMIPALSADDV VIKELKEQLQAKSTGFREKMAFTTPR MMIPALSADDV VIKELKEQLQAKSTGFREKMAFTTPR MMIPALSADDV VIKELKEQLQAKSTGFREKMAFTTPR MMIPALSADDV VIKELKEQLQAKSTGFREKMAFTTPR MMIPALSADDV VIKELKEQLALGNETGRCVVLINGGGG DAASSQEALQAA						ł	
760 2110 A 6077 3 730 PLRLTLMEEVILLJCLKDREGYTSFWNDCISSC LRGCMLIELPRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPLKLHYQLRVKETQPPETVQ NWIELLSGETWNPLKLHYQLRVKETQPPETVQ NWIELLSGETWNPLKLHYQLRVKETQPPETVQ NWIELLSGETWNPLKLHYQLRVKERLAKNIL ALHSDVLENAFAPLLDEDYLATKRVRQLLE LDPEVECLKANTNEVLWAVVAAFTK LALIYL AHASDVLENAFAPLLDEDYLATKRVRQLLE LDPEVECLKANTNEVLWAVVAAFTK HOLD LDPEVECLKANTNEVLWAVVAAFTK HOLD LDPEVECLKANTNEVLWAVVAAFTK HELGINLKAVNQAIQRAELVGEPKNQVTT ACRDAIRSNNINTLEKLWGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKRWNNH TELLGNIKAVNQAIQRAELVGEPKNQVTT ACRDAIRSNNINTLEKLWGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKRWNNH TELLGNIKAVNQAIQRAELVGEPKNQVTT ACRDAIRSNNINTLEKLWGAEDARLMRD MKTMKSRYMELYDLNSSQTDISG SPGINSNNQLAEKVRLRUFEEAKRIANLKI QLAKLDSEAWPOVLDSERLVGEPKNQVT ACRDAIRSNNINTLEKIMRVGTASS SPGINSNNQLAEKVRLRUFEEAKRIANLKI QLAKLDSEAWPOVLDSERVLLINFEKEELLK EMFISPRKVTQGEVEQLEMARKRIEKDLQ ARADTOSKALTERIKLNSKROJLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYTDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITIHEDEVAKTQ KAEGGGRIQALRSLGGFPKSMISLSFEDPELSA TLCELSLGANSQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVVF ASVQRLQASEAAAFDSDESVAGDSGVVF ASVQRLQASEAAAFDSDESVAGDSGVVF ASVQRLQASEAAAFDSDESVAGDSGVVF VSMSYPALHQKTLRVDVCTTORSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGWMAPASGPASTDAVSALLEQTAVE LEKRGGRSTGLEDSWRYEETSNEAVAE EEEEEVEEEEGEEDVFIEKASPDMGGYPALK VDKETNTETPAPSPTVVRKDRXGTPSQGFP LRGSTIIRSKTTSPGPQSQVVCRLNRSDSDSST LSKKPPFVRNSLERRSVRNKKRSPSPPPOPSSVK SLRSSTLISTSLDLELDLQATTWHSQLTQDEIS VLKELKPPLENEVNURGGGG AQISLAEVGQSCKPPPEVGSFREKMAFTTPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPRFSKLVSFNNQESKPEKMAFTTPR MNIPALSADDV VLKELKEQLEQAKSHGEKELPQWLREDERR LLLRMLEKRMDRAEHMGELQTIKMMRAAA KDVHRLRGQGSCKEPPEVGSFREKMAFTTPR MNIPALSADDV VLKELKEQLEQAKSHGEKELPQWLREDERR LLLRMLEKRMDRAEHMGELQTIKMMRAAA KDVHRLRGQGSCKEPPEVGSFREKMAFTTPR MNIPALSADDV VLENLTQGKMCLVPGKTRKLLFKPVAKTED VGKKEITSVOLLALGINETGGRCVVLINVGGGGG DAASSQEALQAARSFKRKLPDNEVHWGSII GAASSQEALQAARSFKRAPLPDNEVHWGSII	ľ	ľ	1	ł			
ILIGOMILELPLRGELQLEACGMRRKSLITRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPILKHYQLRNYRERLAKNI VEKGVLTTEKQNIFLLFPMTTHPI.TINNINIKQR LIKKVQEAULDKWVNDFIMDRRILALIYL AHASDVLENAFAPILDEGYDI.ATKRVRQILE LDPEVECLKANTINEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRFFSTLSVHGLQVDEYHSV MKTMKSRYMELYDLINGYKIRWNNH TELLGNILKAVÑQAIQRAGRILRVGKPKNQVIT ACRAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KRAITCGEKEKQDLIKSLAMILKDGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLILHYEEAKRRIANLKI QLAKILDSEAWPGVLDSSERDRILILINEKEELIK EMFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERIKLINSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSGSSPGSLTSSR GSLVASSLDSSTSASTDLYYDPFEQLDSELQ SKVEFILLEGATGFRPSGCTTHBEDEVAKTQ KAÆGGGRIQALRSLSGFMSTLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLQAV NTAQCGGLKVACVSAAVSDESVAGDSGYYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIIQLSNLSLLQQADQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPSTDAVSALLEQTAVE LEKRQGGRSSTGTLEDSWRYEETSENEAVAE EEEEVEEEGEEGEDVFTEKASPDMDGYPALK VDKENTTETTPAPTVVRKDRXCTPPOQFPY LEKRGGESSTSTJLEDSWRYEETSENEAVAE EEEEVEEEGEGEEOVFTEKASPDMDGYPALK VDKENTTETTPAPTVVRKDRXCTPPOQFSVK SLRSERLIRTSLDLELLOATTWHSQLTQEIS VLKELKGOLQAKSHGEKELPQWLREDERFR LLLRMLEKMBRAEHMGELQTTKMMRAAA KDVHRLRGQSCKEPPEVQSPREKMAFFTRPR MNIPALSADDV VKELKEGLEQAKSHGEKELPQWLREDERFR LLLRMLEKMBRAEHMGELQTTKMRAAAA KDVHRLRGQSCKEPPEVQSPREKMAFFTRPR MNIPALSADDV VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKEITSVOLLALONETGRCVVLNWQGGGG DAASSQEALQAARSFKRAPKLPDNEVHWGSII	760	2110	<u> </u>	6077		720	
VICKSDAPTGDVLLDEALKHVKETQPETVQ NWELLSGETWNPIKLHHYQLRNVRERLAKNI VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR LIKKVQEAVLDKWVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEDYDLATKRVRQLLE LDPEVECLKANTINEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSPHLSGFKFVPRPSFLSVGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAUQRAGRLRVGKPKNQVIT ACRDAIRSNNINITLKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKAMLBOGFKTDRGS SKAITGGEKEKQDLIKAMLBOGFKTDRGS SFGINSNNQLAEKVRLRLRYEEAKRIANLKI QLAKLDSEAWPOULDSEDGRILINEKEELLK EMRFISPRKWTQGEVEQLEMARKALEDDLQ AARDTOSKALTERLKIKNRQLVTRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYDYPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSFRSSLSS PSPPCSPLMADPLAGDAFLNSLEFEDPELSA TLCELSLGNSQRYRLEPGTEGKQLGQAV NTAQGCGLKVACVSAAVEDSEVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQLALK YDEKNKQFAILIQLSNLSALLEQTQQQK VNIR VAVLPCSESTTCLFRITRLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTORSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTILEDSWRYPETSENEAVAE EEFEEVELEEGEEDVTEKASPDMGGYPALK VDKETNTETPAPSTVVRKDRXCOTPSQGPF LRGSTIIRSKTTSDGQSQVVCRLNRSDSDSST LSKKPPFVRNSLERRSVRNKRRSPPPPQPSSVK SLRSERLIRSTLDLELDLGATTWHSQLTQTES LSKRPFVRNSLERRSVRNKKRSSPPPQPSSVK SLRSERLIRSTLDLECLGATTWHSQLTQTES VLKELKEQGSCKEPPEVGTSPCKMAFTTEPR MNIPALSADDV 763 2113 A 6082 3 1558 PPIPRFSKLVSFNNQESYGFCVEREASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKEITSVDLALLGNETGRCVVLINGGGGG DAASSGEALQAARSFKRKLPPNEVHWGSII	/00	2110	^	00//	3	130	
NWIELLSGETWIPLKLIHYQLRNYRERLAKIN VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR LIKKVQEAVLDRWYNDFHRMDRRLALIYL AHASDVLENAFAPLLDEQYDLATKRVRQLLT LDPEVVECKANTINEVLWAVVAAFIK LDPEVVECKANTINEVLWAVVAAFIK LDPEVVECKANTINEVLWAVVAAFIK LOPEVECKANTINEVLWAVVAAFIK HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLINRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRADRISNNINTLFKIMRVGTASS HSDLWSSSSSEPLPKQYLDVSSQTDISG SIGNISNNQLAEVKARLRYEEAKKRIANLKI QLAKLDSEAWPGVLDSSERDRLILINEKEELLK EAMFISPRKWTGGEVELMARRKLEGELQD ARADTOSKALTERLKUNSKRNQLVRELEEAT RQVATLHSQLKSLSSSSSPGSITSSS GSLVASSLDSSTASSTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGLQARSLSGTPKSMTSLSPSSLSS PSPPCSPLMADPLLAGDAFINSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKGFALIQUASVALDQODQKVNIR VANLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVEVEV LEKKQGEGSSTOTILEDSWRYETTSENEAVAE EEEEEVEEEGGEDVTTEKASPDMDGYPALK VDKETNTETPAPSPTVVVRKDDRVCTTPSQOPP LRGSTIIRSKTSPGPQSQYVCRINRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPQPSSVK SLRSERLITSLDLELQATSTHMSQLTQEIS VLKEILKEQLEQAKSHGEKILPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGGSCKEPPEVQSFREKMAFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIFTSKLCVSFNNQEYNOFCVTEEASKANE VLENLTQGKMCLVPGKTRKLPFVAKTED VILENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFVAKTED VLENLTQGKMCLVPGKTRKLPFKVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLPFKVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGKMCLVPGKTRKLLFKFVAKTED VLENLTQGACAGGGGGDAASSQEALQAARSFKRPKLPDNEVHWGSII						ļ	
VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR LIKKVQEAVLDK WVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEQYDLATKRVRQULL LDPEVECLKANTNEVLWAVVAAFTK 161 2111 A 6078 833 390 IVSFILISGFKKFVRFSISVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWINH TELLGNLKAVNQA(GRAGRLRVGKPKNQVIT ACRDAIRSNINITLFKIMRVGTASS 162 2112 A 6079 2 2686 KKATTGEKEKQDLIKSLAMLKOGFRTDRGS HSDLWSSSSSLESSSFPLFKQYLDVSSQTDISG SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI QLAKLDSEAWPGVLOEBEARRRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDFFQLDSELQ SKVEFLLLEGATGFRESPDILLINEKEELLK EMRISPRKWTQGEVEQLEMARRLEEDDELSSL SVEFLLLEGATGFRESPDILLINEKEELLK EARFISPRKWTQGEVEQLEMARRLEEDDELSSL SVEFLLLEGATGFRESPDILLISLEFEDPELSS GSLVASSLDSSTSASFTDLYYDFFQLDSELQ SKVEFLLLEGATGFRESPGLITHEBEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFINSLEFEDPELSS ATLCELSLONSAQERVRLEEPOTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGYTE ASVQRLGASEAAAFDESEAVAATRIQALK YDEKNKQFAILIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYSPAHQKTLTRVPLDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW VSMSYSPAHQKTLTRVPRDASDTLVFNEVFW SLESERLIRTSLDLELDQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKFLPQWLREDERFR LLKMELKRMDRAAHGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRP MNPALSADDV PHPIFFSKLCVSFNIQEYNQFCVIEEASKANE VLENLTQGKCLVPGKTRKLJKKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRPKLPDNEVHWGSII				ĺ			
LIKKVQEAVLDKWYNDPHRMDRRILALIYI							
AHASDVLENAFAPLLDEQYDLATKRVRQLLE LDPEVECKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMEL JUNRDLLNGYKIRWINNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLKKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLKSLAMK, KDGFRTDRGS HSDLWSSSSSLESSSFPLPKQVLDVSSQTDISG SFGINSNNQLAEKVRLRIRYEEAKRRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMOSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYDPFFQLDSELQ SKVEFLLLEGATGGFRSGCTTIHEDEVAKTQ KAEGGGRIQALRSLSGTFKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRILAGSCAAAFDSDESBEAVGATRIQIAK, YDEKNKQPAILIQLSNLSALLQQDQKVNIR VANLPCSESTTCLFTRTPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVQVMAPASPATDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETITTETPASPTVVRPKDRRVGTFSQGFP LRGSTIRSKTFSPGQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPOPSSVK SLRSSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGERKIPQWLREDERRR LLLRMLEKRMDRAFHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PFIFFFSSILCVSFNNQEYNQFCVUEEASKANE VGKKEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRPKLPDNEVHWOSGIG				ł			,
LDPEVECLKANTINEVLWAVVAAFTK		İ		ł			
761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLRSLLVGAEDARLMRD MKTMKSRYMELLVGAEDARLMRD MKTMKSRYMELLVGAEDARLMRD MKTMKSRYMELLVGAEDARLMRD MKTMKSRYMELLVGAEDARLMRD MKTMKSRYMELLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS							
HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSFPLFKQYLDVSSQTDISG SFGINSNNQLAEKVRLIRYEGAKRRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMRISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSMØGLSSGSSFGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSCGTTHHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQCGGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESBAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQDQKVNIR VAVLPCSESTITCLFRIRPLDASDTLVFNEVFW VSMSVPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGRSTRYWNLLSYKYLKGG RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGGEDVTEKASPDMGGYPALK VDKETNTETPAPSPTVVRKDRRVGTPSQGPF LRGSTIIRSKTTSPGPQSQVCRLNRSDSDSST LSKKPPFVRNSLERSVVMKRFSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLERKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRKLFNDNEVHWGSII							LDPEVECLKANTNEVLWAVVAAFTK
MŘÍTMKSRYMELVDLINGYKIRWINH TELLGINLKAVNQAIQRAGRILRVGKPKNQVIT ACRDAIRSININITLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSLESSEFPLKQYLDVSQTDISG SFGINSINQLAEKVRLRIKYELAKRINLKI QLAKLDSEAWPGVLDSERDRILIINEKEELLK EMRFISPRKWTQGEVEQLEMARKRIANLKI QLAKLDSEAWPGVLDSERDRILIINEKEELLK EMRFISPRKWTQGEVEQLEMARKRIEKDLQ AARDTQSKALTERLKLINSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRSPGCTITHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TILCELSLGNSAQERVRLEEPGTEGKQLGQAV NTAQGGGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIQLONLSASLLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEDVFTEKASPDMDGYPALK VDKETINTETPAPSPTVVRKDRRVGTPSQGPP LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVBMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKEI,PQWLREDEFISP LLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRKLPDNEVHWGSII	761	2111	Α	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSPLPKQYLDVSSQTDISG SFGINSNQLAEKVRLRLRYEEAKRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMOSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFINSLEFEDPELSA TLCELSLGNSAQERVRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIQUSNLSALLQQDQXVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMGGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPP LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVNSLERRSVRMKRSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLIRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		ľ					HQKLSADMADHSNLIRSLLVGAEDARLMRD
762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSJESSSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLRYEEAKRRIANLKI QLAKLDSEAWFGVLDBERDRLILDEKEELLL EMRISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKINSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSSSPSGLITSS GSLVASSLDSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFINSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTPLDASDTLVFNEVFW VSMSYPALHOKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSPHEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETTNTETPAPSPTVVRPKDRRVGTPSQGFF LRGSTITIGSKTFSPGPQSQVYCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MIPALSADQ MIPALSADG T63 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRPKLIPNEVHWGSII	}	l		ł			MKTMKSRYMELYDLNRDLLNGYKIRWNNH -
762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSPLPKQYYLDVSSQTDISC SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMFISPRKWTQGEVQLEMARKRLEKDLQ AARDTQSKALTERLKINSKRNQLVRELEEAT RQVATLHSQLKSLSSSMSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS		!		l			TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
HSDLWSSSSLESSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLRLRYBEAKRRIANIKI QLAKLDSEAWPGVLDSERDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRIQATGFRSSMGSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRAQARSLSGTPKSMMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NITAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFFTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGGSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTTSPGPQSQVVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCEPEVQSFREKMAFFTRPR MIPPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNGFCVIEEASKANE VLENLTQGKMCLUPGKTRKLIFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		İ	ļ.				ACRDAIRSNNINTLFKIMRVGTASS
SFGINSNNQLAEKVRLRLYEEAKRRIANI.KI QLAKLDSEAWPGVLDSERDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIRQLSNLSALLQQQQK VNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPP LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNGFCVIEEASKANE VLENLTQGKMCLUPGKTRKLLFKFVAKTED VGKKJEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII	762	2112	Α	6079	2	2686	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS
SFGINSNNQLAEKVRLRLYEEAKRRIANI.KI QLAKLDSEAWPGVLDSERDRILLINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIRQLSNLSALLQQQQK VNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPP LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNGFCVIEEASKANE VLENLTQGKMCLUPGKTRKLLFKFVAKTED VGKKJEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		İ	[4		HSDLWSSSSSLESSSFPLPKOYLDVSSOTDISG
QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLINSKRNOLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS REILKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEGEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDEFFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MIPPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFRNQEYNGFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPJMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGGEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKFLPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII			ł				
AARDTQSKALTĒRLKLNSKRNQLVRELEEĀT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFILLEGATGFRPSGCITTHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFALIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVPNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVWAPASGPASTDAVSALLEQTAVVE LEKRQEGRSSTOTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPFFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKLEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWQSII	-						
RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCITTHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPAAHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEGEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		İ					
GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFPRSGCITTHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGGEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKFLPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKLEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRPKLPDNEVHWGSII							ROVATI HSOLKSI SSSMOSI SSGSSPGSI TSSR
SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTILEDSWRYETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKFLPQWLREDEFFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQG RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPPPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKPRSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		}	ŀ				
TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII				1			
NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAÆ EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKEI.PQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRQGSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII				.			
VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTILEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII				[
AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		[
RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII		-					
LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII			1	1			
EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPOPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							RELKPVGVMAPASGPASTDAVSALLEQTAVE
VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							EEEEEVEEEGEEDVFTEKASPDMDGYPALK
LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF
LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST
SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
1 LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV 763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII]		, I	
KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV							
763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII					ļ		
763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII						•	, ,
VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII	763	2112		6000	-	1660	
VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII	103	2113	А	0082	٥	1008	
DAASSQEALQAARSFKRRPKLPDNEVHWGSII							
I I I I I I I I I I I I I I I I I I I		1					
					i		TO LOW ALCOHOLD DOLLAR THOUSE A CONTROL

فيه فيها فيه فيدر است الراسد الله الله

086 18	Lancin	1 57 -	Loro	I B . I		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucicc	J	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 40,700			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
}				sequence		nucleotide insertion
						VIVQSHEKTQIRDVKLTAGLKPGQDANLTQK
			l			THVTLHGTELCDESYPALLTDIPVGDLHPGEQ
			İ	 		LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE
	İ	ĺ	ì	}		KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF
		1				EHLERVYADIPFLLMTDLLSASPWALTIVSSE
l		İ				LHLAPSMTTVDQLESQVDNVILQTGESASECF
		ĺ	1			CLQCPSLGNIEGGVATGHYIISWKRTSAMENI
j		l	1	•		PITTVITLPHVIVENIPLHVNADLPSFGRVRES
ŀ		}				LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG
		1				LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
!]			LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
74	0114		6000			DDTSIAAA
764	2114	Α	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP
						PPPAPAPPTLGNNHQESPGWRCCRPTLRERN
		1				ALMFNNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA
		ĺ				AFLILLKYMYSDEIDLEADTVLATLYAAKKYI
						VPALAKACVNFLETSLEAKNACVLLSOSRLF
		!				EEPELTQRCWEVIDAQAEMALRSEGFCEIDR
1		1				QTLEIIVTREALNTKEAVVFEAVLNWAEAEC
ļ		}	ł			KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE
		l				FANGAAQSDILTLEETHSIFLWYTATNKPRLD
						FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG
 						RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV
]		•				KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF
·						EHPVQVEQDTFYTASAVLDGSELSYFGQEGM
			}			TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE
						LIFYA
765	2115	Α .	6099	1	1150	SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF
			i i			RPVKAPGTFHMVHGKCMCKHNTAGSHCQH
					•	CAPLYNDRPWEAADGKTGAPNECRTCKCNG
						HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGOVCOBCVBGEVBDI BBBESABDACVBCS
						TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA
						GRRCDRCMVGYWGFGDYGCRPCDCAGSCD
						PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP
						WEWEDAQGFSALLHSGKCECKEQTLGNAKA
						FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK
						KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL
						NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK
						PSLGRKVMDILKRECK
766	2116	A	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR
						CVLTERGLQLFEAKGTGGRPKELSFARIKAVE
						CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW
						NAQITLGLVKFKNQQAIQTVRARQSLGTGTL
						VS
767	2117	A	6106	i	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP
						MARYYIIKYADQKALYTRDGQLLVGDPVAD
				İ		NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC
						LACVETEEGPSLQLEDVNIEELYKGGEEATRF
						TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ
2.0						PVQLTKESEPSARTKFYFEQSW
768	2118	Α	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF
						YTPSVISSVMHRVARCAAPHVHILLANFYLLF
260	2110		L			PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	Α	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS
						SSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ
						ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY
	L				L	KHEDLQTDESSMDDRHPRRQLCGGNQAATE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RIILFGRELQALSEQLGREYGKNLAHTEMLQD AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
770	2120	A	6125	2	570	NSAILESQNLPKQPPLMLALGQASECLRLMA RAGLGSCSFARVDDYLH YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
						VAPWALKYMNRRASQMLLMFLLAICLLAIIF VPQEMQMLREVLATLGLGASALANTLAFAH GNEVIPTIIRARAMGINATFANIAGALAPLMM ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	A	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC LTKEDTGWYWCGIQRDFARDDMDFTELIVT DDKGTLANDFWSGKDLSGNKTRSCKAPKVV RKADRSRTSILIICILITGLGIISVISHLTKRRS QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	А	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV WYLLRKHWIANNLFGLAFSLNGVELLHLNN VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV ALAKGEVTEMFSYEESNPKDPAAVTESKEGT EASASKGLEKKEK
773	2123	A	6161	3	1088	CQPMLVTRKNHPKLLLRRTESVAEKMLTNW FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV NPENENAPEVPVKGLDCDTGTQAKEKLLDA AYKGVPYSQRPKAADMDLEWRQGRMARIIL QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH LDQREGDRGSKMVSEIYLTRLLATKGTLQKF VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ ADKHQIHDADVRHTWKSNCLPLRFWVNVIK NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA GSEGAGLPPSGELHFWVKEARDLLPLRAGSL DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF NHTMVYDGFGPADLRQACAELSLWDHGALA NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT DMKKSPEIISRRMTFAL*CYSLTFVRFAHYVQ VPWNWLMLGCHTAVDFDQLISSMPCISHGMT ASASAL
776	2126	A	6217		827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT RQKHAKKHLGFFRNNFGVREPYQILLDGTFC QAALRGRIQLREQLPRYLMGETQLCTTRCVL KELETLGKDLYGAKLIAQKCQVRNCPHFKNA VSGSECLLSMVEEGNPHHYFVATQDQNLSVK VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA VESGRLSQCMRKKVSNISKRNRV**KTLNRG RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE KKRKRKRIRNRSNPKVLSEKQNAEGE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence 1038	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL* FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ RFORGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/ SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC NSFRYRR
779	2129	Α	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP YGQSQPSCFDRVKMGFVMGCAVGMAAGAL FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH QSQPMY
780	2130		6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG TDYWLYSRGVCRTKSTSDNETSRKNEEVMT HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT PV
781	2131	A .	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH LQVKQKQLACLCTWQARDPDCPPSTKVVL/L VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD LEFLGDLKGCSELKNFQELITQSALVHPKADV WWYCGRPLLGTLPSN
782	2132	Α	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD DDKKRVKAKKKKKKKKKKKKKKKKKKKKK ESSDSSCKDSEEDLSEATWMEQPNVADTMDL IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM SGSRHRRMEAVRLRKENQIYSADEKRALASF NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRREAAEGLHFLGPPGRVRGQ LRGITGPAWYCHSPSHSI.I.SAFCHLPTPSRCP AMARPPVPGSVVVPNWHES/RRGQGVPGLHS AQEPPAGVWAA*AASAAAAILSIDTASYKIFV SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ TTVVFWPAKLQASSRVVMFEFFWDCGESA LKKFDHMILLACMENTDAFLFLFSFTDRASFE DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA PRKPEGAQARTAQSGALRDVSEELSRQLEDIL STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RSKLESLCRELQRHNRSLKEEGVQRAREEEE KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR QENMELAERLKKLIEQYELREHIDKVFKHK DLQQQLVDAKLQQAQEMLKEAEERHQREKD FLLKEAVESQRMCELMKQQETHLKQQLALY TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI KKLEKETTMYRSRWESSNKALLEMAEEKTV RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
785	2135	A	6319	1493	889	GQRWGSHRTSAVRIFS SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT KLPWSWGMRPMKIFFSEEYRSISTRISHDAL* EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVRKNENCYNTP GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMTV KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	Α			6622	PRSLCFSLWAEAAVLADGGLRRRRLLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/VWRRDQRPERRELL*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNILKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

SEQ ID NO: of nucle peptide sequence No: of nucle otide sequence No: of nucle otide sequence No: of nucle otide sequence Sequence No: of nucle otide sequence Se	C=Cysteine.
nucle cotide sequence uence Deptide sequence USSN Ocation	
eotide sequence USSN 09/496 corresponding 1 location on the form of the first amino acid residue of peptide sequence Poptide sequence Residue of peptide of pe	
sequence Sequence	
uence 914 ap to first amino acid residue of peptide sequence peptide sequence 914 sequence 915 sequence 916 sequence 917 Fireonine, V=Valine, W=Tryp Y=Tyrosine, X=Unknown, **Stop-possible nucleotide deletion, Y=Tyrosine, X=Unknown, **Stop-possible nucleotide deletion, Y=Tyrosine, X=Unknown, **Stop-possible nucleotide deletion, Y=Tyrosine, X=Valine, W=Tryp Y=Tyrosine, X=Unknown, **Stop-possible nucleotide deletion, Y=Tyrosine, X=Valine, W=Tyryp X=Tyrotione, Y=Tyrot	
amino acid residue of peptide sequence y-Valine, W-Tryorsine, X-Unknown, *-Sto /-possible nucleotide deletion, *-possible nucl	
residue of peptide sequence	
peptide sequence incleotide deletion, Vanceotide insertion RQNSEREAGKKHK VEDGTSS' MSLFSPSIKQDAPRTTSHARP VSYTDLDNI, PNSDEDELTPOS ASCKESKTGNLDPLSCISTADL EQHIMGFSPMNMNNKEYGSG GNSSIGAQFKEVDEGFCSPK KPENCQILVGCSMFAPLKTLP' CTYRQSWTVGKLELLSSGPSM DDEYGTAYTPCHITSCGMPP LPSPSTRFFTPRTPRTTPRTTPRTPRTP VSYTDLSHAMPORTORITSCOMPP LPSPSTRFFTPRTPRTPRTPRTPRTPRTP VX VENSDLYSPSATSTCRPLN EAHSLYVNLILSESVMNLFKD NMNIKGADVGYYPDDTPGAC MRKFGNNSGLFFEDELDIGG KRFEALRATSAEHVNGGLKE QDOCTNLFSPFGAADODFFK EERDCCNDCYLALEHGRQFM ALVKSSCLIPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGW KMAGRGSVGTDESPEPLPIPTT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFREDLTAIVESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLLYAQVCRVDL GNNEAFSKLKLLYAQVCRVDL SLLSQPNLVAPTSQSLTTPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLTOPHPWSKGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLTOPHPWSKGVSSNKLPSI AGSMSTQANTVQSGGLGGQQ ESSSLTOPHPWSKGVSSNKLPSI AGSMSTQANTVQSGGLGGQQ ESSSLTOPHPWSKGTSTMORD AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTY PVKHEDRETYPPQLKSLAFSAI TNVKTLTGFOPGLAMETALRS PPFILAPVKDKQTLGFTFGEA GYCLSHDQRWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKWE LPWRVVIGRLGRIGHGELKWE SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTTQDTSCTHILVPTTSASVQV DLAFNFNNDGADGMGIFDLI MLPASFTGSFVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCDLALG LPDWFWSACPQAQVCQLALG LPDWFWSACPQAQVCQLSL	
sequence nucleotide insertion	
RQNSEREAGKIHK VEDGTSS' MSLFSPSIKQDAPRPTSHARPP VSYTDLDNLFNSDEDELTGS ASCKESKTGNLDPLSCISTADL EQHMGFSPMINNINKEYGSM GNSSIGAQFKIEVDEGFCSPK KFENCQL VGCSMFAPLKTLY CYPROSWTVGKLELSSGFSM DQEYGTAYTPOTHTSCGMPPS LPSPSTPRFPTRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRT	ossible
MŠLFSPSIKQDAPRPTSHARPP VSYTDLDNLFNSEDELITPGS ASCKESKTGNLDPLSCISTADL EQHIMGFSPMNMNKKEYGSM GNSSSIGAQFKIEVDEGFCSPK KPENCQILVGCSMFAPLKTLP: CTYRQSWTVGKLELLSSGFSM DQEYGTAYTPQTHITSCGMPPS LPSPSTPRPPTPRTPRTPRTPRTPRC VKYENSDLYSPASTISTICPLI EAHSLYVILLISESVMILFKD MMIKGADVGYYIPDPTQEAG MNRKFGNNSGLFFEDELDIIGI KRFEALRATASAHVNGGLKEE QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGW KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLINGAKSFFROLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SILSQPNLVAPTSGYLTIPPQM ATLASAASSTMTVTSGVAISTS AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTD GLLEFLEMVATTYCSTALSSA TNVKTLTGFGPGLAMETALRS PPFILAPVADKQTELGETTGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSAKFGLQKLWE PPFILAPVADKQTELGETTGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSAKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRRGGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNDGADGMGIPDLLI NILPASPTGSPVHSPGSAMPS	
VSYTDLDNLFNSDEDELTGS ASCKESKTGNLDPLSCISTADI EQHMGRSPMMNNKEYGSM GNSSIGAQFKIEVDEGFCSPK KPENCQILVGCSMFAPLKTLP: CTYRQSWTVGKLELLSSGPSM DQEYGTAYTPQTHTSCGMPP LPSPSTPREPTPRTPRTPRTPRTPR VK YENSDLYSPASTPSTCRPLD EAHSLYVILILSESVMILFKD NMNIKGADVGYIPDPTQEAG MRKFGNNSGLFPEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAJQKKTVRPWGW KMAGRGSYGTDESPEPLPIPTT LSPFALPYWERLMLEPYGSQR NEALLINGAKSFREDLTAIYESS LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SILLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMRDDI AVTYPPAIVVYIIDPFTYENTD GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARRFGLOKLWE LPWRVVIGRLGRIGHGELKDM SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVPPTSASVQV DLAFNPNNDGADGMGIDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLB	
ASCKESKTÖNLDPLSCISTADI EQHIMGFSPMNMNKEYGSM GNSSSIGAOPKIEVDEGFCSPK KPENCQIL VGCSMFAPLKTLP. CTYRQSWTVGKLELLSSGPSM DQEYGTAYTPQTHITSCGMPPS LPSPSTPRPTTRTPRTPREDE VKYENSDLYSPASTPSTCRPLN EAHSLYVLILLSESVMNLFKD NMNIKGADVGVYIPDPTQEAC MMRKFGNNSGLFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGV KMAGRGSYGTDESPEPLPIPTI LSPPALPYWERLMLEPYGSQR NEALLNGAKSFFDLTAIYESQ LLTDGIMRVGSTASKKLSKLL GNNEAFSLKLYAQVCRYDL SLLSQPNLVAPTGSSTTSVKLSKL GNNEAFSLKLYAQVCRYDL SLLSQPNLVAPTGSVSTTSVANS AGSMSTQANTVGSQLGGQQ ESSSLPTQPHPDVSESTMDRD AVTYPPAIVVYIIDPFTYENTD GLLRCFLEMVQTLPPHIKSTVS PVKHEDRETYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETTGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVPPTSASVQV DLAFNPNDGADADGMGIDLLI NILPASPTGSPYNSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	STSLIYDSDLA
EQHIMGFSPMMNNKEYGSM GNSSIGAQFKIEVDEGFCSPK KPENCQIL VGCSMFAPLKTLP CTYRQSWTVGKLELLSSGPSM DQEYGTAYTPQTHTTSCMPPS LPSPSTFRFPTPTTRTTRTTRTPRC VKYENSDLYSPASTPSTCRPLN EAHSLYVNLLISESVMNLFKD NMNIKGADVGVYIPDPTQEAC MNRKFGNNSGLFEEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAJQKKRTVRPWGV KMAGRGSYGTDESPEPLPIPTI LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESC LLTDGIMRVOSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSINNSGVSSNKLPSI AGSMSTQANTVQSQQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPALVYVIIDPFTYENTD GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TINVKTLTGFOFGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSSVUMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSSVHSRGSNYPHG RILSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCLFLG LPDWFWSACPQAQVQCLFLG LPDWFWSACPQAQVQCLFLG	KRSANGSDDK
GNSSSIGAOFKIEVDEGFCSPK KPENCQILVGCSMFAPLKTLPY CTYROSWTVGKLELLSSGFSM DQEYGTAYTPQTHTSCGMPPS LPSPSTRFPTPTTTTTFTRTFOR VKYENSDLYSPASTFSTCRPLN EAHSLYVNLILSESVMNLFKD NMNIKGADVGVYIPDPTQEAG MRKFGNNSGLFFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDFFPK EERDCCNDCYLALEHGROFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYDTESPFEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESS LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTOPHPDVSESTMDRDF AVTYPPAIVVYIIDPFTYENTDI GGLRCFLEMVQTLPHIKSTVS PVKHEDREFYPQHLKSLAFSAI TNVKTLTGFOPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHORWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMGGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSSVHSPGSNYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLB QSDELLHSKHSHPLDSNQTSD	HKMYPTPPSL
KEENCQILVGCSMFAPLKTILP CTYRQSWTVGKLELLSSGPSM DQEYGTAXTPQTHITSCGMPP LPSPSTRRFPTRTRTPRTPRTPR VKYENSDLYSPASTFSTCRPLI EAHSLYVNILLSESVMNLFKD NMNIKGADVGVYIPDPTQEAG MRKFGNNSGLFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHFWSKRNDVSMQ LQPVLQDAJQKKRTVRPWGVG KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDJ AVTYPPALVYJIIDPFTYENTD GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLSLAFSAI TNVKTLTGFOFGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGS NRARRKSSAKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMGISADDSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSSVHSRGSNYPHG RILSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLB USDWFWSACPQAQQYQCPLFLB QSDELLHSKHSHPLDSNQTSD	DTTPGGTVLE
CTYRQSWTVGKLELLSSGPSM DQEYGTAYTPQTHTSCGMPPS LPSPSTTPRFPTRPTRTPRTRTRG VKYENSDLYSPASTPSTCRPLN EAHSLYVNLLISSEVMNLFKN NMNIKGADVGVYPDPTQEAG MNRKFGNNSGLFFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGW KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL- SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSILNSGVSSNKLPSI AGSMSTQANTVQSQQLGGQQ ESSSLFTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDRETYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKGTELGETFGEA GYCLSHDQRWILASCTDLYGG NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPINDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCPLFLK QSDELLHSKHSHPLDSNQTSD	PSEIKDFSYVY
CTYRQSWTVGKLELLSSGPSM DQEYGTAYTPQTHTSCGMPPS LPSPSTTPRFPTRPTRTPRTRTRG VKYENSDLYSPASTPSTCRPLN EAHSLYNLLISSEVMNLFKN NMNIKGADVGVYPDPTQEAG MNRKFGNNSGLFFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL- SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSILNSGVSSNKLPSI AGSMSTQANTVQSQGLGGQQ ESSSLFTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDRETYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKGTELGETFGEA GYCLSHDQRWILASCTDLYGG NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCPLFLK QSDELLHSKHSHPLDSNQTSD	GYLPLIKLPEE
DQEYGTAYTPQTHTSCGMPPS LPSPSTPRFPTPRTPRTPRTPRTPR LPSPSTPRFPTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPRTPR	-
LPSSTRFPTPRTPRTPRTPROV VKYENSDLYSPASTPSTCPUP EAHSLYVNLILSESVMNLFKD NMNIKGADVGVYIPDPTQEAG MNRKFGNNSGLFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDFFFK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVG KMAGRGSYGTDESPEPLPIPTI LSPFALPYWERLMLEPYGSQR NEALLINGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCCRYDLI SLLSQPNLAYDTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNINSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPFAIVYYIIDPFTYENTDI GLLRCFLEMVQTLPPHKSTVS PVKHEDRETYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNDGADGMGIFDLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCPLFLB QSDELLHSKHSHPLDSNQTSD	
VKYENSDLYSPASTPSTCRPLN EAHSLYYNLILISESVMNILFAD NMMIKGADVGYVIPPDTQEAG MNRKFGNNSGLFFEDELDIIGI KRFEALRATSAEHVNGGIKES QDQCTNLFSPFGAADQDFFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESC LLTDGIMRVGSTASKKLSEKL GNEAFSKLKLYAQVCRYDL. SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMVTSGVSSNKLPSI ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLIRCFLEMVQTLPPHKSTVS PVKHEDRETYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQR WILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVIGGLGRIGHGEILDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD	
EAHSLYVNLILSESVMNIFKD NMMIKGADVGYPIPDTQEAG MNRKFGNNSGLFFEDELDIIGI KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGROFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPT LSPFALPYWERLMLEPYGSQR NEALLINGAKSFRRDLTAIYESS LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDL SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAIST ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSQQLGGQQ ESSSLPTQPHPDVSESTMDRD9 AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGF NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNFNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPWWSACPQAQYQCPLFLB USDBELLHSKHSHPLDSNQTSD	-
MMNIKGADVGVYIPDPTQEAC MMRKFGNNSGLFEEDELDIIGI KRFEALRATSAEHVNIGGLKES QDQCTNLFSPFGAADQDFFFK EERDCCNDCYLALEHGRQFM AL VKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPTF LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAFTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSQQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREITYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PFFLLAPVKDKQTLEGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLI QSDELLHSKHSHPLDSNQTSD'	
MNRKFGNNSGLFFEDEDIGIG KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPTF LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSINNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVYVIIDPFTYENTDI GLLRCFLEMVQTLPPHKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLI QSDELLHSKHSHPLDSNQTSD	
KRFEALRATSAEHVNGGLKES QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVG KMAGRGSYGTDESPEPLPIPTF LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASFTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLI QSDELLHSKHSHPLDSNQTSD'	
QDQCTNLFSPFGAADQDPFPK EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPTF LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLITDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDF AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETTGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCPLFLB QSDELLHSKHSHPLDSNQTSD	
EERDCCNDCYLALEHGRQFM ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPLPIPTE LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLITDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDF AVTYPPAIVYJIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALSA PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSF EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQVQCPLFLI QSDELLHSKHSHPLDSNQTSD'	
ALVKSSCLHPWSKRNDVSMQ LQPVLQDAIQKKRTVRPWGVV KMAGRGSYGTDESPEPPIPITE LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDIG AVTYPPAIVVYIIDPFTYENTDIG GLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAITNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEAGGYCLSHDQRWILASCTDLYGG NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLIG NILPASFTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLIG LPDWFWSACPQAQQQCPLFLIG LPDWFWSACPQAQAQQCPLFLIG LPDWFWSACPQAQAQQCPLFLIG LPDWFWSACPQAQAQCPLIFL LPDWFWSACPQAQAQCPLIFL LPDWFWSACPA	
LQPVLQDAIQKKRTVRPWGVO KMAGRGSYGTDESPEPLPIPTE LSPFALPYWERLMLEFYGSQR NEALLNGAKSFFRDLTAIYESQ LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLO SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDE AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE USDELLHSKHSHPLDSNQTSD'	
KMAGRGSYGTDESPEPLPIPTT LSPFALPYWERLMLEPYGSQR NEALLNGAKSFFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDF AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETTGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD'	•
LSPFALPYWERLMLEPYGSQR NEALLNGAKSFRDLTAIYESG LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDIS AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD'	
NEALLNGAKSFFRDLTAIYESC LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLA SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDIA AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
LLTDGIMRVGSTASKKLSEKL GNNEAFSKLKLYAQVCRYDLG SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDG AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TINVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RILSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
GNNEAFSKLKLYAQVCRYDLO SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	-
SLLSQPNLVAPTSQSLITPPQM ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDI AVTYPPAIVVYIIDPFTYENTDI GLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE' LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD'	-
ATLASAASSTMTVTSGVAISTS ASTSSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDE AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TINVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
ASTSSSSSNLNSGVSSNKLPSI AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDR AVTYPPAIVVYIIDPFTYENTDI GLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TINVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLR QSDELLHSKHSHPLDSNQTSD	
AGSMSTQANTVQSGQLGGQQ ESSSLPTQPHPDVSESTMDRDM AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLM QSDELLHSKHSHPLDSNQTSD	
ESSSLPTQPHPDVSESTMDRDM AVTYPPAIVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RILSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLM QSDELLHSKHSHPLDSNQTSD	
AVTYPPAJVVYIIDPFTYENTDI GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TINVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD	
GLLRCFLEMVQTLPPHIKSTVS PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
PVKHEDREIYPQHLKSLAFSAI TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	ESTNSSSVWTL
TNVKTLTGFGPGLAMETALRS PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
PPFILAPVKDKQTELGETFGEA GYCLSHDQRWILASCTDLYGE NRARRKKSSARKFGLQKLWE' LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD'	TQCRRPLPTS
GYCLSHDQRWÏLASCTDLYGE NRARRKKSSARKFGLQKLWE' LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD'	PDRPECIRLYA
NRARRKKSSARKFGLQKLWE' LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD'	GQKYNVLFV
LPWRVVIGRLGRIGHGELKDW SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLI NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLk QSDELLHSKHSHPLDSNQTSD	LLETCINIDVP
SLSKRLKDMCRMCGISAADSP EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	VCLGLVQMSS
EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	SCLLSRRNLQ
EPQGSFVIMPDSVSTGSVFGRS NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	
NTPQDTSCTHILVFPTSASVQV DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	TTLNMQTSQL
DLAFNPNNDGADGMGIFDLLE NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLE QSDELLHSKHSHPLDSNQTSD	ASATYTTENL
NILPASPTGSPVHSPGSHYPHG RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLK QSDELLHSKHSHPLDSNQTSD	
RLLSTEPHEEVPNILQQPLALG LPDWFWSACPQAQYQCPLFLI QSDELLHSKHSHPLDSNQTSD	
LPDWFWSACPQAQYQCPLFLI QSDELLHSKHSHPLDSNQTSD	•
QSDELLHSKHSHPLDSNQTSD	
MNML	
789 2139 A 6359 1 2002 TGTLTEDGLDVMGVVPLKGQ	AEL DI VIDEDDE
LPVGPLLRALATCHALSRLQD	
1 1 1 1 1 1 1	
MVESTGWVLEEEPAADSAFGT	
LWEPQLQAMEEPPVPVSVLHR	•
SVVVAWPGATQPEAYVKGSPI	
VPTDFAQMLQSYTAAGYRVV	
SLEAAQQLTRDTVEGDLSLLG	
QTTPVIQALRRTRIRAVMVTGI	
RGCGMVAPQEHLIIVHATHPE	•
LSGPTFGIIVKHPPKLLPKVLVC)GTVFARMAP
EQKTELVCELQKLQYCVGMC	3DGANDCGAL
KAADVGISLSQAEASVVSPFTS	SMASIECVPM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR NITDTGFKLLLVGLVTLNFVGGLHAGERARP VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW
790	2140	Α	6380	76	1059	PPLPAGPLR SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCFRGFEVMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/LEKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALGYFHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN
791	2141	A	6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VTCPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQEPIYDLTSVPNLFTY\SS FNN\GQLAPGIT\MTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK
792	2142	A	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL
793	2143	Α	6446		152	PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL LEYCMVTGYDWWDILLHVQPSMVQSLVEKL HEEYTRQTAALQQVLSTRILAMKASLCKLSP CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN QPCPTSEPCFTSEPSPTSEPSPTSEPSSP*SLCG SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC RDEGPASEPDEALVDECCLLPSQLLIPSLDWL PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT LQLDGLARAPGQPKIDHLRRHLGACPTEEC KACTRCGCVTMLKSPNRTTAVKQWEQRWIK NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT PRSLDHLHPEDRP
794	2144	Α	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG GSIEPRDLRLQ*AVITPLYTPAWVTQ
795	2145	Α	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM KKMYKLQNISEELQRNISLQLMSNMNISNKIR NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR WIWHKDSCYFLSDDVQTWQESKMACAAQN ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE DSYYSWYESG*YNQVPSAWVIRNAPDLNNMY CGYINRLYVQYYHCTYKQRMICEKMANPVQ LGSTYFREA
797	2147	A .	6507	1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH APRAASAQLEERMRDPHPGMTLQEGDCRGS QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA AIWTGAVAVLAGAAAFIYEKRGGTYWALLR TLLALAAFSTAIAALKLWNEDFRYGYSYYNS ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM DMLKALFRTLQAMILGVWILLLLASLTPLWL /SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG STHASARSQVPRSAGEAAPHSRRPPGLLPHAP RAASAQLEERMRDPHPGMTLQEGDCRGSQT VSLTMGTADSDEMAPEAPQHTHIDVHIHQES ALAKLLLTCCSALRPRATQARGSSRLLVASW VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI WTGAVAVLAGAAAFIYEKRGGTYWALLRTL LALAAFSTAIAALKLWNEDFRYGYSYYNSAC RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM LKALFRTLQAMLLGVWILLLASLTPLWLYC WRMFPTKGVSP
798	2148	A	6528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP RFLVAFAYWNHYLSCTSPCSCYRPLCKLNFG LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

SEQ D SEQ D No. of nucle order Predicted end Dr. of nucle order Predicted Predicted end Dr. of nucle order Predicted P
nuciotide cotide contesponding of the contesponding of peptide residue of peptide sequence of peptide
Institute Inst
corresponding no last amino no last mino no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino scid residue no last mino
unince 914 anj to first anino acid residue of peptide residue of peptide sequence peptide sequence
amino acid residue of peptide sequence
Part
Poptide sequence
EVPPEGTIGLPLPHISDLPTSWCGHISLQCGSQSS FPPAHEMAPIPASSIGHMLQCGSQSS FPPAHEMAPIPASSIGHMLQCGSQSS FPPAHEMAPIPASSIGHMLQCGSQSS FPPAHEMAPIPHASSIGHMLQCGSQSS FPPAHEMAPIPHASSIGHMLQCGSQSS GNASRQVGRNSVEGLUGLGIGNKLRVV GQNLGL*HCVWVWETGE*KWRLQMGEP GVASRRQVFNSVEGLUCHNSSAPPMYMGFF SPTVPGGGVGG-HLVTFLHIPPEVEAAGIPLLL GPSLPQRQGEIIIVVILAAPACAPFIDR*WEP RERPSP*ELGLGEPTLSYPAFHATAWWF RERPSP*ELGLGEPTLSYPAFHATAWWF GNKELLTISQPEEKRP FPFQRINFIEHSGSVSLLALACDLGWCEDWS CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV DEARCKESQQEAGGERQRGARTHHWRGW EKGRRVRLRPPSKGLEADQPVRKLGGPTPSYT ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVEPBTRGLCEERFACKHL QHIEGSERRHEETTRQAALDGEPLGGGQLTA VHLHPSKEQQGQEGGERQRGARTHHWRGW EKGRRVRLRPPSKGLRADQPVRKLGGPTPSYT ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVEPBTRGLCEERFACKPEPRPPPL GLPGDPTGPVTHHAPPVSTGASGQERRAEP GAYSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGGRLLALLLLVPGPGGAS EITTELPDNAKOCFYEDLAGGTCLEFGVTTG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCSNEUSSGGREPTEGSM GEALLLLVVSIGQVFLKSFSSTFTHKTVYTEDPQVG ETHLCFLVRDRVSALTGMESACVSHEALKS VIDYQTHFRLREAQGRSRAEDLNTRAYWSV GEALLLLVSIGQVFLKSFSSDKRTTTTRVG GRALLLLVSIGQVFLKSFSSDKRTTTTRVG TASKNGTYKFCSNEUSSGREPTEGSM DGWRMPRWGUVFLKSFSDKRTTTTRVG SGLALLLLVSIGQVFLLSKSFSDKRTTTTRVG HKLFABSDSSSVHMGTELTLRYGTOVSGGL VKSTYYNRDSBRVHGTGELTRYGTOVSGGL SQLALLLVSIGQNFPISTONRSGVKSCSSLYTACVY HKLFABSDSSSYHMGTELTLRYGTOVSGGL SQCALLLVSIGGSSPCHYEE GREPTYTYRDSBRVSLAGGSSPCHYEE GREPTYTHRDSBRVSLAGENSTUCKED VFSTYYNRDSBRVSKLISTLAHAPIPPTGPTT SADYLFQESVSSKKLSTLAHAPIPPTGPTT SADYLFQESVSSKKLSTLAHAPIPPTGPTT LAGATTRIKFYTIEFDRGNPHGFALAR KRIFLKRMSIRESLKERGDRONPHGFALAR KRIFLKRMSIRESLKERGRONPOPVNGMVHVIKG UPGSQIKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFISTNICHNHYIG UPGFSQIKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFISTNICHNHYIG UPGFSQIKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFISTNICHNHYIG UPGFSQUKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFISTNICHNHYIG UPGFSQUKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFISTNICHNHYIG UPGGSQUKYSCTKGYRLIGSSSATCLISGDTVIW DEFTPICARICGCIPPITTNOFIGTNICHNHYIG UPGGSVUKYSCTKGYRLIGSSSATCLISGDTVIW
FPPAHIENAFIVFIASSLGAMENCILLIKITKK HTVSQEDGISLAGAPRQPRRKSRTSVLRIRV MVRWELSSNGNPGROVLGLGLGGNKLRVV GQNLGL*HCVWVVWETGP*KRWRLQMGE# GVASRRQ*VRNSVRGLVCINSSAPPMYMGF SPTVFGGVGG*LHVTTILHIPPEVEAAGPILL GESLPQRQGREIHIVVILAAPACAPPHDR*WEP RERPSP*ELGLRGEPTLSYPASCRVIRQPIPD RKSYSWKQRLINFISFFSALAVYFRHNMYC EAGVYTIFAILEYTVVLTNMAFHMTAWWDF GNRELLITSOPEKRF GNRELLITSOPEKRF CCLVQGGGILVVUTNIGEDEAGGDTDSV DEARCKESQGEAQENLREDLCLSFAKDKIL QUIEGSERRHEETTTKQAALDGEPLGGQLTA VHLHPSKEQQGGGGGRQRAFATHWRGW EKGRRVRLRPPSGKLRADQPVRKLIGGTTPST ELPGLQPHAPTPTTAPATTYSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTRKACPFKPRPL GLQPPATPTTTAPATPTTPPPV RWKCPLPVEPRTRQLCRERTRKACPFKPRPL GLQPPATPTTTAPATPTTPPPV RWKCPLPVEPRTRQLCRERTRKACPFKPRPL GLQPPATPTTTAPATPTTPPPV GHYDVDCRLEDPDGKVLYKEMKQYDSTTT TASKNOTYKFCPSNESSTITHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSHFEALKS VIDYOTHFEILEAGGRSRAEDRTGTST TASKNOTYKFCPSNESSTITHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSHFEALKS VIDYOTHFEILEAGGRSRAEDRFTGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIELKRMPSIRESLERGYDMARLGPFBVSQP MKALTLGNTTSSVILTNYMDTQYYGEIGITT PROFINCAGE DGVAGMGFEQGAGRAPPOTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIELKRMPSIRESLERGYDMARLGPFBVSQP MKALTLGNTTSSVILTNYMDTQYYGEIGITT PQTFKVVPDTGGSTSVILTNYTGTVSGVIL GFSHYTNILKTGGVVLKB GFSHYTNILKTGVGVJCMSCSTLCE DGCLALVDTGASYSGSTSSIEKLMPALGAKE KRIFDYVVKCNEGFTIPTT-FLLGGKDTPTT SADYLFGESSSKKLSTLAHAMYIPPTGPTI VLGATNIRKFYTEDRGNOPHGFALAR KRIFDYVKCNEGFTIPTT-FLLGGKOTPTT VLGARMGASSPSSFEFVGPFGCGGSDC VFSFYYNRDSENSGSLGGOVLGGSSPQUFE GFSHYNLIKTGVWQLOMK GVSSTILCE DGCLALVDTGASYSGSTSSIEKLMPALGAKE KRIFDYVKCNEGFTIPTT-FLLGGKOTPTT VLGARMGASSPSSFEFVGPFSGICLKNS VWTGAKDRCRRSSCRNPDPVNGMVHVKG UPGSQUKVSCTKGYRLJGSSSATCISGDTVIW DWETTCICRIPCGLPFTITNODFISTNRENFHY UPGTFCORPGCUFFNKTDFTVINDGISCDTVIW DWETTCICRIPCGLIFTNTNODFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGERSITCTSND DQVGWSGPAPQLIFNKCTPTNYLGLYSND DQVGWSGPAPQLIFNKCTPTNYLGLYSND DQVGWSGPAPQLIFNKCTPTNYLGLYSND DQVGWSGPAPQLIFNKCTPTVNYLGLYSND DQVGWSGPAPQLIFNKCTPTVNYLGLYSND DQVGWSGPAPQLIF
HTYSQEDGISLAGAPRQPRKKSRTSYLRIV MVRWELSSNORPGROVLGIGISMLRVY GQNLGI-HCVWVVWETGFYKBWRLQMGEB GVASRQVRNSVRGLVCHNSSAPPMYMGFF SPTVFGGGVGG*LHVTFILIPFEVEAAGIPLLL GFSLFQRQREIIIVVILLAAPACAPFIDIR*WEP RERPSP*BLGLAGEPILSYPASRVIRQPIP*D RKSYSWKQRLFIINFISFSALAVYFRHNMYC EAGVYTIFALLEYTVLTNMAFHATTAWWDF GNKELLTISQPEEKRF 799 2149 A 6529 1 874 FPFPGRINIERSGSVSLLAACDLGWCEDWS CCLVQGGGDLVDVVQTNIGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QHIEGSERHEETRTKQAALDGEPLGGGGLTSV UNIHPSKEQQQEGGERQRGARTHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGFTPST ELPGLQPHAPTTPHTAPATFTYSPAPDTPNPPV RWKCPLPVEPSTRQLCREERKACPFPKPPPL GLPGDPTGPVTHHAPPVSPTGASGQERAEP GAYSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGGRLLALLLLVFGPGGAS EITTELPDNAKQCFYEDLAGGGS EITTELPDNAKQCFYEDLAGGGS EITTELPDNAKQCFYEDLAGGGS EITTELPDNAKQCFYEDLAGGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLVKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFTHKTYYFDFQVG ETTHLCFLVRDRVSALTQMESACVSHBALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALLLLVVSIGQVFLKSFSDKRTTTTRVGS ETTHLCFLVRDRVSALTQMESACVSHBALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALLLLVVSIGQVFLKSFSDKRTTTTRVGS BOUNGRIPHPTTAPTTAPTYCYGEIGGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFASDSSSYHMGTELTLTRYSTGTVSGGI, SQDIITYGGTTYTQMFGGVTEMPALPFMLAEP DGVVGMGFTEQAGRVTPIFDNILGGVVKED VSSFYYNRDSENSSGLGGJVIGGSDPGPTE GRFHYNLIKTOVWQCHAGVSVSCSSLLCE DGCLALVDTGASYNGSTSSISKLMEALGAKE KRLFDYVVKCNEGFTLPTFT-FLLGGKDTPLT SADYLFQESVSSKKLSTLAHANJPPTTGTT. LACATTRIKFYTTEPDRONPHGFALAR SQDIITYGGTYVMCCREGFTGPFTDTTT SADYLFQESVSSKKLSTLAHANJPPTTGTT. LACATTRIKFYTTEPDRONPHGFALAR SQDIITYGGTYVMCCREGFTGPFTTNLTPTTGSGTLLCEN UNTGAKADRCRRRSCENPPDPVNGMVHVIKG UPFSFPIGTATNYCCRPGVGFRIGCTGSGI LAVVVLLALPVAWGGCNAPEWILPFARFTNL TDEFFPIGTATNYCCRPGVGFRIGCTSSI LAVVVLLALPVAWGGCNAPEWILPFARFTNL TDEFFPIGTATNYCCRPGVGFRIGCTSSI LAVVVLLALPVAWGGCNAPEWILPFARFTNL TDEFFPIGTATNYCCRPGVGFRIGCTSSI LAVVVLLALPVAWGGCNAPEWILPFARFTNL TDEFFPIGTATNYCCRPGVGFRIGCTSND DQVGWSGPAPQCIPPKCTPNYLGGNUNDLUSD UPGTFCORRICCGLPHTTNODFTSTNEENHY UPGGSVLYYSCNYGTHANGELVSD DQVGWSGPAPQCIPPKCTPNYLGGNUNDLUSD UPGGWSGPAPQCIPPKCTPNYLGGNUNDLUSD UPGGWSGPAPQCIPPKCTPNYLGGNUNDLUSD UPGGWSGPAPQCIPPKCTPNYLGGNUNDLUSD
MYRWELSSNGNPGROLGLGLGLGINKLRVV GQNLG1-HCVWVVWETGRWRLVMGFE GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF SPTVFGGVGG*LHVTFILHPFEVSAGIPLLI GFSLQRQGG*G*LHVTFILHPFEVSAGIPLLI GFSLQRQGRGHINVILAPACAPPHDR*WEP RERFSP*ELGLRGEPTLSYPASCRVRQPFPD RESYSWKQRLFINFISFSTSALAVYFRHNMYC EAGYVTIFALLEYTVVLTMAFHINTAWUDF GNKELLITSQFEKRF RESYSWKQRLFINFISFSTSALAVYFRHNMYC EAGYVTIFALLEYTVVLTMAFHINTAWUDF GNKELLITSQFEKRF GNKELLITSQFEKRF CLLVQGGDLVDVVQTNHGEDEAGGDTDSV CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV CLLVQGGGDLVDVVQTNHGEDEAGGDTDSV CLLVQGGGDLFEDLCLESFAKDKIL QIIEGSBREHEETRTKQAALDGEPLGGGQLTA VHLHPSKEQGQGGGGGRGARIHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGFTPST ELPGLQPHAPPTHTM7APTTYSPADTTNPPV RWKCPLPVEFRTRQLCRERTRKACPPKPRPPL GLPGDPTIGPYTHAPPVSPTGASSGGRRAEP GAVSYAHASATK GHOPVDCRLEDPDGKVLYKEMKQCPSFTT TASKNGTYKFCSSNESTENTKVYPEDFQGAS ETTELPDNAKQCFYEDIAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKQVDSFTF TASKNGTYKFCSSNESTENTKVYPEDFQVG ETHILCFLVRDRVSALTQMESACVSHEALKS VIDYQTHERHEAGQGRSKDEDITRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTRVGS GEALILLVSIGQVFLLKSFSDKRTTTRVGS GEALILLVSIGQVFLLKSFSDKRTTTRVGS GEALILLVSIGQVFLLKSFSDKRTTTRVGS GEALILLVSIGQVFLLKSFSDKRTTTRVGS SQDITVGGITVTQMFGEVTEMPALPFMLAEP DGVVGMGFEQAIGK VPTSTNIGQVFLKED VPSFYVNRDSENSQSLGGQIVLGGSDPQHFE GNFHYNILLKTGVWOJGMKSVGSSTLLCE DGCVAGMGFEQAIGK VPTSNINGQVLKED VPSFYVNRDSENSQSLGGQIVLGGSDPQHFE GNFHYNILLKTGWWOJGMKSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDVVKCNBGPILIPTFIFLLGKDFTPLT SADYLFQESYSSKLISTALHAMYIPPTTGPTL VLGATFURKFYTETDRGNPHIGFALAR KRLFDVVVKCNBGPTLPFTFIFLAGKENTPLFTPTIFLLGKOTPTLFT SADYLFQESYSSKLISTALHAMYIPPTTGPTL VLGATFURKFYTETDRGNPHIGFALAR VMCLGRMGASSPSFSPEPOFVNGMVHVKG QFGSQIKYSCTKGYRLIGSSATCIISGDTVIW DNETPICDRPCCIPPTTNNGDFISTRREMPHY GSVYTYRCHPGSGGRKVEELVGERSYCTSND DQVGIWSGPAPQCIIPNKCTPNVENGLVSD DVGGSQIKYSCTKGYPRLOVINGD
GQNLGL*HCVWVVETGE*KRWRLQMGIE* GVASRRQ*VRBNYGRLYKRSAPPMYMGFF SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL GFSLPQRQGREIHVVILAAPACAPFHDR*WEP REFRS*PELGRIGHEPTVVLTAPACAPFHDR*WEP REFRS*PELGRIGHEPTVVLTAPACAPFHDR*WEP RESYSWKQRLFINFISFSALAVYFRHNMYC EAGVYTTRAILEYTVVLTAMAFHMTAWWDF GNKELLTISQPEEKRF 799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLGWCEDWS CCLVQGGGDLVDVQTNHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFADKIL QIRGSSREHEBETRTKQAALDGEPLGGGQLTA VHLHPSKEQQGOEGGERQRGARTHWRGW EKGRFVRLRPFSGKLRADQPYRKLGGPTPSYT ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVERTRQLCRERTKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEPQVTTG GHYDVDCRLEDPDGKV, YKEMKKQYDSFTT TASKNGTYKFCFSNEFSTFTHKTVYFDFQVG ETHLCFLVRORVSALTOMACVSHEALKS VIDYQTHFRLREAGGRRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSIKRTITTRVGS VIDYQTHFRLREAGGRRAEDLNTRVAYWSV GEALILLVSIGGVFLLSSFSIKRTITTRVGS DGWRMPRWGLLLLLWGSCTFGLPTDTTFF KRIFLKRMPSIRESLKERGVSGREFQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVSGREFQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVSGSREFTHACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGTTVTJGMFGEVTISNGVVKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHTE GRYPHYNLIKTGVWQIQMAGLAGPBWSQP MKR.TLGNITSSVILTNYMDDTQYYGEIGIGTP PQTFKVVFDDTNISGVVKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHTE GRYPHYNLIKTGWVQIQMKGVSGSTLLCE DGCLALVDTGASYSGSTSSIEKLMBALGAKE KRLPDVVVKCNEOPTLPFLLGGKDTPLT SADYLFQESYSKKLSTLAIHAMYTPPTGPTL VALGATPIRKFYTETDRGNNPHIGFALAR 802 2152 A 6567 13 6147 MCLGRIMGASSPRSPEPVGFPPFCLGGSL LAVVVLLALPVAWQCCNAPEWLPFARPTHL TDFEFFFGTINNYECRGPSPTCISKNS VWTGAKDRCRKSCRNPPPDPVNGMVHVKG (1QFGSQUKYSCTKGYRLLIGSSAATCIISGDTVIW DNETPICDRIPCQLPPTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVEELVGERSYCTSND DQWGIWSGPAPQCIIPNKCTPPNVENGLVSD DQWGIWSGPAPQCIIPNKCTPNVENGLVSD DQWGIWSGPAPQCIIPNKCTPNVENGLVSD DQWGIWSGPAPQCIIPNKCTPNVENGLVSD DQWGIWSGPAPQCIIPNKCTPNVENGLVSD
GVASRRQ*VRNSVRGLVCHNSSAPPWYMGFF SPTVFGGGVG*LHVTFILHPPEVBAAGPILL GPSLPQRQGG*LHVTFILHPEVBYBAAGPILL GPSLPQRQGGEIHVVILAAPACAPPHDR*WEP REIRSP*ELGLRGEPTLSYFASCRVRQPIPD RKSYSWKQRLPINFISFSTSALAVYFRHNMYC EAGVYTIFAILEYTVVLTMAFHMTAWWDF GNKELLTISQPEKRF 799 2149 A 6529 1 874 FFFFQRINFIEHSGSVSLLALACDLGWCEDWS CCLVQGGDLVDVQTNHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIEGSEREHEETRTKQAALDGEPLGGGQLTA VHLHPSKCQQGGGGGGRAGENGERTHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGPTPST ELPGLQPHAPTPHTAPATPTYSPAPDTPNPV RWKCPLPVEPRTRQLCRESFRKHACPPKPRPPL GLPGLOPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGFGAS ETHELPLVARACCYSEDAGCTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFTHKTVTFDFQVG ETHLCFLVKDRVSALTQMESACVSHEALKS VIDYQTHFRLREAGGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSPSKRTTTTRVGS GEALILLVSIGQVFLLKSPSKRTTTTRVGS GEALILLVSIGQVFLLKSPSKRTTTTRVGS GEALILLVSIGQVFLLKSPSKRTTTTRVG MKR.TLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTDTSSNVWYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGVSGFL SQDJITVGGITVTQMFGGVTEMPALPFMLAEF DCVVGMGFIEQAIGRVTPIFDNISQQVLKED VFSFYYNRDSENSSLSQUQVLCGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDVVVKCNEGPTLPTFTFLLGGKDTPLT SADYLPQSSYSSKLSTLAHHAMYPPPTGPTL VALGATFIKFYTEFDGROMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDVVVKCNEGPTLPTFTFLLGGKDTPLT SADYLPQSSYSSKLSTLAHHAMYPPPTGPTL VALGATFIKFYTEFDGROMWHVIKG 10FGSQUKYSCTKGYRLIGSSATCIISGDTVIW DNETPICDRIPCQLPPTTNGDFISTNRENFHY GSVYTYKCNPGSGRKVFELVGERSYCTSND DQVGIWSGPAPQCIIPNKCTPNVENGLVSD DQVGIWSGPAPQCIIPNKCTPNVENGLVSD DQVGIWSGPAPQCIIPNKCTPNVENGLVSD
SPTYFGGYGG-LHVTFILHPPEVEAAGPILLI GPSLOPRGCBILIVULACAPHDRE-WEP REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D RKSYSWKQRLFIINFISFFSALAVYFRHNMYC EAGVYTTEALLEYTVVLTNMAFHMTAWWDF GNKELLITSQPEEKRF FPFGRINPIEISGSVSLLALACDLGWCEDWS CCL-VQGGGDL-VDVVQTINHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIEGSSREHEETRTKQAALDGEPLGGGQLTA VHLHPSKEQQGGGGGRQRGARTHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGFTPSVT ELPGLQPHAPTPHTAPATPTYSPAPDTINPPY RWKCPLPVEPRTRQLCRERKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCPYEDIAQGTKCTLEPQVTTG GHYDVDCCLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFHKTVYFDFQVG ETHLCFL VKRORVSALTOMESACVSHEALKS VIDYQTHFELREAGGRSRAEDLNTRVAYWSV GEALILLVVSIQQVFLLKSFFSDKRTTTTRVGS SGEALILLVSIQQVFLLKSFFSDKRTTTTRVGS WKRILLAWSGCTGLIPTTTT KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRITLGNTTSSVILTNYMDTQYYGGIGGTP PQTFKVVFDTOSSNV WYSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGTVTYGNFFORMFLETARAY FDGWVGMGFIEQAIGRVTPIFDNINSQCVLKED VRSFYYNRDSENSQSLGGQVUGGSDPQHYE GNFHYNLIKTGVWQQIGMSVSGSTLLCE DGCLALVDTGASYSIGSTSSIEKLMBALGAKE KRIPDVVVCKNEGPTLPFTLFLLGGKDTPLT SADYLPGESYSSKLSTLAHHAMYPPPTGFTL VALGATFIRKFYTEFDRONNPHIGTALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTHL TDEFEFFIGTYLNYECRGPSPTSICLKNS VWTGAKDRCRRKSCRNPPDPVNGMHVIKG IQFGSQKYTSCTKGYBLLOSSSATCTISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVEELVGERSYCTSND DQWGWSGSTKGYBLLISSSATCTISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVEELVGERSYCTSND DQWGWSGFRAPCHIPNKTENFLYCTSND DQWGWSGFRAPCHIPNKTENFLYCTSND DQWGWSGFRAPCHIPNKTENFLYCTSND DQWGWSGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND DQWGWGFRAPCHIPNKTENFLYCTSND D
GPSLPQRQGEBIIVVILAAPACAPPHDR*WIPP REIRPSP*ELGLRGEPTY-SASCRVIRQPIPD RKSYSWKQRLFIINFISFFSALAVYFRHMYC EAGYYTIFAILEYTVVLTNMAFHMTAW WDF GNKELTISQPEEKBP GNKELTISQPEEKBP CCLVQGGGDLDVVVOHGEDEAGGDTDSV CCLVQGGGDLDVVVOHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIIEGSEREHEETRTKQAALDGEPLGGQCTA A HILPSEKQQQEGGERQRGARTHHWRGW EKGRRVRLRPPSGKLRAQDPVRRLGGPTPSIT ELPGLQPHAFTPHTAATTYSPADTPNPPV RWKCPLPVEPRTROLCRERTRKACPPKPRPPL GLPQDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFVEDAGTKCTLEFQVITG GHYDVDCRLEDPDGKVLVKEMKKQYDSFTF TASKNGTYKFCFSNEFSFTFTHKTVYFPDFQVG ETHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHERLREAQGRSRAEDLINTRVAYWS SUDYQTHERLREAQGRSRAEDLINTRVAYWS GEALILLVSIGQVPLLKSFFSDKRTTTTRVGS GFALILLVSIGQVPLLKSFFSDKRTTTTRVGS GFALILLVSIGQVPLLKSFFSDKRTTTTRVGS MKRTLGNTTSSVILTNYMDTGYGEIGGTV PQTFKVVFDTGSSNV WYPSSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDITVGGTVTGMFGGTWALPFFFLILGGKDFPLT SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYSSKLSTLAIHAMYIPPFTGPTL SADVLFQESYGRFGICKLKNS VWTGAKDRCRRRSCRNPPDPVNGMVHVIKG IQFGSQUKYSCTKGYRJAGSSSATCIISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGGGGRAVFELVCGEPSIYCTSND DQVGIWSPAPQCLIPPNKCTPPNVENGILVSD
REIRPSPYELGLRGEPTLSYPASCRVIRQPIPPD RKSYSWKQRLFIINFISFSTALAVYFRHNMYC EAGYYTIFAILEYTV-LYTIMAFHMTAW WDF GNKELLTISQPEEKBF FFFFQRNPIEHSGSVSLLALACDLGWCEDWS CCLVQGGGDLVDVVQTNHGEDEAGODTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIEGSEREHEETRTKQAALDGEPLGGGQLTA VHLHPSKEQQQGGGGRQRGARTHHWRGW EKGRVRIRPPSKIRAOPVPKLIGOFTPSIT ELPGLQPHAPTPHTAPATTYSPAPDTPNPPV RWKCPLPVEPRTRQLCREFTRKACPPKPPPL GLYGDYTHHAPPVSPTGASGQERRAEP GAVSYAHASATK SOO 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFVEDLAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKKQYDSTIT TASKNGTYKFCSSNEVESTITHKTVYTDFQVG EITHLCFLVR/DRVSALTQMESAACVSHEALKS VIDYQTHFRIREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS VIDYQTHFRIREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS WRKITLGHTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGITVTQMFGEVTEMPALPFMLAEF DGVVCMGFIEQAIGRVTPIFDNISQGVLKED VFSSFYYNTDSENSOSLGOVLGGSDVLGGDPHYE GNFHYNLIKTGVWQQMKGVSVGSSTLLCE DGCLALVDTGASYSISGSTSSIEKLMEALGAKE KRIPDYVVKCNEGFTLPPTFIFLIGGKDTPLT SADYLPQESYSSKLSTLAIHAMYIPPPTGPTL LAUGATTPIKFYTTETEGRGVTPTPV GNFHYNLIKTGVWQQMKGVSVGSSTLLCE DGCLALVDTGASYSGSTSSIEKLMEALGAKE KRIPDYVVKCNEGFTLPPTFIFLIGGKDTPLT SADYLPQESYSSKLSTLAIHAMYIPPPTGPTL LAUGATTPIKFYTTETGRORPHFYALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPYGPPACLFPCCGGSL LAVVVLLALPVAWQCNAPEWLPFARPINL TDEFEFFIGTYLNYECREGYSGRFSICLKNS VWTGAKDRCRRKSCRNPPDPVNIGMYHVIKG 10FGSQIGVSCTKGVGLIGSSSATCLISGDTVIW DNETPICDRIPCGLPPTTNOFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCILIPNKCTPPNVENGILVSD
RKSYSWKQRLFINFISFFSALAVYRRHNMYC EAGYVTYEALEYTVVLTNMAFHMTAWWDF GNKELLITSQPEEKF 799 2149 A 6529 1 874 FFFFQRINFIEHSGSYSLLALACDLGWCEDWS CCLVQGGGLVDVVQTNHGEDEAGGDTDSV DEARCKESQQBAQENLREDLCLESPAKDKLL QIEGSREHEETRIKQAALDGEPLGGGQLTA VHLHPSKEQQGGGGERQRGARTHHWRGW EKGRVRLRPPSGKLRAQPVKRLGGFTPS/T ELPGLQPHAFTHTHATATFTYSPADTINPPV RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDDLAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLVKEMKKQYDSFTF TASSKIGTYKFCFSNEVSTHKTVYYFDFQVG ETHLCFLVR/DRVSALTQMESACVSHHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECKGEGLREPQNLSGSGREPQTEGSM MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWYPSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDITVCGTTVTCMTGTTFTFLTLGGKOTPLTF GRIFYNTDSENSQSLGVTLGSDSTOPLYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSIERLMEALGAKE KRIFDTYVKCCHGGTVTCMTGGTVTGMTGGTVTGMTGGTTPTFTLFLLGGKOTPLT SADVLPGESYSSKLISTLAHAMYIPPTGTTL SADVLPGESYSSKLISTLAHAMYIPPTGTTL LUGATFIRKFYTETEDGNTHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWQCCNAPEWLPFARPINL TDEFEFFIGTYLNYECRERGYSGRFSIELKNS VWTGAKDRCRRSCRNPPDPVNGMYHVIKG IQFGSQUKYSCTKGVRLIGSSSATCUSGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGGPSIYCTSND DQVGIWSPAPQCLIPNKCTPPNVENGLLVSD
BAGVYTIFÁILEYTVVLTNMAFHMTAWWDF GNKELLITSQPEEKRF FFFFQRINFIEHISGSVSILALACDLGWCEDWS CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV DEARCKESQQEAGNLREDLCLESFAKDKIL QIEGSEREHETRITKQAALDGEPLGGQLTA VHHPSKCESQQEAGNERDLCLESFAKDKIL QIEGSEREHETRITKQAALDGEPLGGGQLTA VHHPSKCESQQEAGGRATHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGPTBST ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV RWCPLPVEPRTROLCERETRKACPKPRPPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK GASYAYATKA GAVSYAHASATK GAVSYAHASATK GASYAYATKA GAVSYAHASATK GASYAYATKA
GMKELLITSOPEEKRF 799 2149 A 6529 I 874 PFFFQRINFIEHSGSVSLLALACDLGWCEDWS CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QHEGSERHEETRIAQALDGEPLGGGQLTA VHLHPSKEQQGQEGGERQGARTHHWRGW EKGRRVRLRPSGKLRADQPVRKLGGPTPS/T ELPGLQPHAPTPHTAAPTTYSPAPDTDNPPV RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLALLLUVPGPGGAS EITHELPDMAKQCPYEDIAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFTHKTYYTDFQVG ETHLCFLVRDRVSALTQMESACVSHEALKS VIDVQTHERLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 I 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKGYDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTOTYYGEIGIGTP PQTTKVVFDTOGSSNVWYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIRVFIPDNINGQVLKED VFSFYYNRDSENSQSLGQQIVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGSVGSSTLLCE DGCLALVDTGASYJSGSTSSIEKLMRALGAKE KRLPDYVVKCNEGPTLPFTFLLGGKDTPLT SADYLFGESYSKKLSTLAHAMYIPPPTGPTL ALGATPIKRFYTEPDRSNPHGFALAR 802 2152 A 6567 I3 6147 MCLGRMGASSPRSPEPVGFPAGLPFCCGGSL LAVVVLALPVAWQCCAPFEVLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRFSIICLKNS VWTGAKDRCRKSCRNPPDVNGMYHVIKG IQFGSQKYSCTKGYRLJGSSSATCLISGDTVIW DNETPICDRIPGGLPPTTTNODFISTNRENFHY GSVVTYRCNPGSGGRKVSCTKGYELVGEPSIYCTSND DQVGIWSGPAPCLIPINKCTPPNVENGILVSD
PFFPQRINFIEHSGSVSLLALACDLGWCEDWS
CCLVQGGGDLNDVVQTINHGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESFAKDKIL QIIEGSEREHEITRTKQAALDGEPLGGQLTA VILHPSKEQQGQEGGERGARTHHWRGW EKGRYVRLRPPSGERTKACPPKRLGGPTPS/T ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV RWSCPLPVEPRTRQLCRERTKACPPKRPPPL GLPGDPTGPVTHHAPPVSPTGASGQERAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITTELPDNAKQCFVEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGRVLYKEMKKQVDSTTF TASKNGTYKPCFSNEVSTFTHKTVYFDPQVG EVTHLCFLVR/DRVSALTQMESACVSHEALKS VIDYQTHFRLREAQGRSRAEDLNTXVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 I 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMFSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAGIGRVTPIFDNISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHTE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGFILPSTFTLFLLGKGEDTPLT SADYLFQESYSSKLISTAHAMYPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSFEPVGPPAGLFFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPINL TDEFEFPIGTYLNYECRPGYSGRPFSICLKNS VWTGAKDRCRKKSCRNPPDPVNGMVHVIKG QSVYTVRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGAPQCIIPNKCTPPNVENGILVSD
QIJEGSEREHETRTKQAALDGEPLGGQQLTA VHLHPSKEQQGQEGGERQRGARTHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASQGERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITTELPDNAKQCFVEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQVDSFTF TASKNGTYKPCFSNEVSTFTHKTVYFDFQVQ EVTHLCFLVR/DRVSALTQMESACVSHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDJITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYNNLKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSIEKLMEALGAKE KRLFDYVVKCNEGPTLPFTFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGFTL VALGATFIRKFYTEETDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPEPVQPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFFIGTYLNVECRFGYSGRPSSICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETFICDRIPCGLPPTTINDGPTSTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGRAPQCIIPNKCTIPNVENGILVSD
Vilhipskeqqqogegerqraakithungw Ekgrevrlrppsgerradqcvrklggptps/T
EKGRRVRI.RPPSGKLRADQPVRKLIGGPTPS/T ELPGLQPHAPTPHTAPATTYTSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRILLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSTTI TASKNGTYKFCFSNEFSTSTHKTVYFDFQVG EXTHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHFRIREAQGGRSRAEDLNTRVAY'NSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 I 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSISELKERGYDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VPSFYYNRDSENSQSLGGQIVLGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTIF-LLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATPIRKFYTEETPGNNPHIGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEPFIGTYLNYECGREGSSATCIISGDTVIW DNETPICDRIPCGLPPTITNOEDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGOPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFTHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHERLREAQGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 I 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGITP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGFILPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAHHAMYIPPPTGFTL VALGATFIRKFYTEFDRGNNPHGFALAR R02 2152 A 6567 I3 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPSIICLKNS VWTGAKDRACKSCRNPPDPVNGMVHVKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL GLPGDPTGPVTIHAPPVSPTGASGQERRAEP GAVSYAHASATK 800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEEQVITG GHYDVDCREDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEVSTFTHKTVYFDFQVG ETHLCFLVRDRVSALTQMESACVSIHEALKS VIDYQTHERLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGFTLPFTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAHAMYIPPPTGPTL VALGATPIRKFYTEFDGRNPHGFALAR KRLFDYVVKCNEGFTLPFTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAHAMYIPPPTGPTL VALGATPIRKFYTEFDGRNPHGFALAR VWTGAKDRASSFSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNVECRFGYSGRPFSILCLKNS VWTGAKDRACKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPTTITNGDFISTNRENFHY GSVVTYRCNPGSGGKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
GLPGDPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK
800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEVFSTFTHKTVYYEDFQVG EITHLCFLVR/DRVSALTQMESACVSIHEALKS VIDVQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYNILIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLPDYVVKCNEGPTLPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL ALGATTPIKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFFEPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMYHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGRPFTITTNODFISTNRENFHY GSVYTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITTELPDNAKQCFYEDIAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG E\THLCFLVR\DRVSALTQMESACVSIHEALKS VIDYQTHFREAQGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLUWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRL\TACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGTTVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAHAMYIPPPTGPTL ALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICRIPGCLPPTITNOEPTISTNRENFHY GSVYTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
BITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNEFSTFTHKTVYFDFQVG ETHLLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS B01 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLIPFARPTNL TDFEEPPIGTYLNVECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVYTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTTYRFCFSNEFSTFTHKTVYFDFQVG E\THLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHERLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREP/DLSGSGREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GFFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPFTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAHMAMYIPPTTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVULALPVAWGQCNAPEWLPFARPTNL TDEFEFFIGTYLNYGCRGYSGRFFSICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
TASKNGTYKFCFSNE/STFTHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TIDEFEFFIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGVRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
801 2151 A 6556 I 1319 TPCMECIKGGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VPSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAIHAMYIPPPTGPTL VALGATF/IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLWGSCTTGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLUWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLLSTLAIHAMYIPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLFFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
801 2151 A 6556 I 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 I3 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPSIICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSISEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKLISTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRFFSIICLKNS VWTGAKDRCRKSCRNPPDVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GFFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFFIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
DĞVVGMGFIEQAİGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLİKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAİHAMYIPPPTGPTL VALGATF\İRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFİSTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
GNFHYINLIKTGVWQIQMKGVSVGSSTILCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\RKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
KRLFDYVVKCNEGPTLPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\IRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
LAVVVLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
NRSLESI NEVVEER COPGEVMK GPRR VK COA
LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS TNRENFHYGSVVTYRCNPGSGGRKVFELVGE PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP RRVKCQALNKWEPELPSCSRVCQPPPDVLHA ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG RHTGKPLEVFFFGKAVNYTCDPHPDRGTSFD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
	·					PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEQFFFASFTIPINDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEFFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP
803	2153	A	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

MSERVSGLAGSIVREFERLIURYDEEVVELLE LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKRERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGVKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKRNELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSSLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIDAVQPGDILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVDIS QTGYYTEHVPTDPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLDDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ							· · · · · · · · · · · · · · · · · · ·
nucleotide seq- uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 uence USSN 09496 1914 Imminio acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide Sequence Sequence Sequence DSSUBLE ALL VIGAVIETLA HGG YVTD GENT QRMW NESPQUEVIELLA HGG YVTD GENT GENT GENT GENT GENT GENT GENT GENT	-		1				
uence USSN genee 99496 corresponding to last amino acid residue of peptide sequence properties amino acid residue of peptide sequence peptid			поа	1	, , ,		
sequence 99496 corresponding on the first of the private incidence of peptide residue of peptide sequence computer sequ					L		, , , , , , , , , , , , , , , , , , , ,
uence 1914 main aeid main aeid residue of peptide sequence T-Thronine, V-Valine, W-Typtophan, Y-Tyrosine, V-Valine, W-Typtophan, Y-Tyrosine, X-Unknown, *-Slop coden, P-possible unclotide delicin, *-possible muclotide disention DGILLY/QAVETLA-HGGTV/DGENT/QRNW NESPCGVTGV/LIDSSGDRETDFSLWDMDPG NGARVAU/NYNGTSQLE/UAVSGRELAN/PLG YPPPDIPKCGFDNEDPACNODHLSTLEVIAL/W GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELASE, WAV GSLSLLGILU/SPFTYRRMQLERELSTLD/WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV DGRYU-KITDYGLESTRID-WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV GGRY-FRENCHSILU/SPFTYRMGDL-PRESILD/WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV GGRY-FRENCHSILU/SPFTYRMGDL-PRESILD/WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV GGRY-FRENCHSILU/SPFTYRMGDL-PRESILD/WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV GGRY-FRENCHSILU/SPFTYRMGDL-PRESILD/WMFRY SLTNDIVKGMLF-IHNGALGSHGNLKSSNCVV GRAND-PRESILE GRAND		1 -	ĺ	1			
## ## ## ## ## ## ## ## ## ## ## ## ##	•	uence	ĺ	l .			
residue of pepide sequence	uence	[ļ	914			
## peptide sequence ## peptide sequence ## peptide sequence ## peptide servion ## pep			1	i			
Sequence nucleotide insertion		Ì	l	l		sequence	y=1 yrosine, X=Unknown, *=Stop codon,
DGILLYQAYTETLAHGGTYTDGENTQRMW NRSFQGYTGYLRIDSGORRETPSELWYMDDE NGAFRVVLNYNGTSQELVAVSGRKLNWPLG YPPPDPKCGFONEDGENCHONGUSTLEVLALV GSLSLLGILVSFYYRKMQLEKELASELWRVW WEDVEPSSLERHLRASGRITLSGGSNYGSI LITTEGQFQVPAXTAYYKGNLVAVKRVNRXR IELTRKVLFELKINRDVQNBHLTREVGACTD PPPNCLTTYCPGGSLQDILENESITLDWMFPY SITNDIVKGMLELINGALGSHONLKSSNCVV DGEFVLKITDVGLESFRDLDPEQGHTVYAKK LVTAPELRMASPYRGSQAGDVYSGGILQE IALRSGVFHYEGLD.SKEIERVTYRGEQPFRF PSLALGSHLBELGLLMQRCWAEDPQGRPFFQ QIRLTLRKPRRENSSNLDNLLSRREQYANNL EELYBERTQAYLEERKARALALLYQILPHSVAE QLKRGETYQABAPDSVTTYFSDIVGTALSAE STPMQVVTLINDLYTCFDAVIDNEDVTKVET IGDAYMWVSGLPWRGNLHACEVARMALAL LDAYRSFRRRPQGQLRINGHTGHTGPVCAGV VGLKMPRYCCPGDTVATSRNEGGALMI HLSSETKAYLEERGGFPQTVATSRNGGALMI HLSSETKAYLEERGGFPQTQCAMAAAA UVALENLONDATSRNGGALMI HLSSETKAYLEERGGFPQTQGAMAAAA LDAYNGGALTAYSTAAAT AGRETAYATAGAAATATATATATATATATATATATATATATATA		}]			
NRSFQGVTGYLKIDSSGDRETDFSLWDMDDE NGAFRVUNTNYGTGSGLVAVSGRKLNWPLG YPPPDPKCGFONEDPACNODHLSTLEVLALV GSLSLLGHLVSFTYKKMQLKEKLASEL WRVW WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSS LTTEGGPOVPAKTAYKGNLVAVKRVNRKR IELTRKVLJERLKHMRDVQNBHLTREVGACTD PPNICLTEYCPFGSLDDLENESTILDWMFRY SLTNDIVKGMLFLHNGAICSHONLKSSNCUV DGEPVLKITDYGLESFEDLDPSGOHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGILQE HALRSGVFTVGLDENESTILDWMFRY SLTNDIVKGMLFLHNGAICSHONLKSSNCUV DGEPVLKITDYGLESFEDLDPSGOPERFFG QIRLTRKRNRENSSNILDMLISRMEQYANNL EELVEERTQAYLEERKRAEALLYQHJPHSVAE QUKRGETVQAEAFDSVTTYFSDIVGFTALSAE STFMQVVTLINDLYTCFDAVIDNPDVKVET IGDAYMVVSGLPVRNGRLHAGEVARMALAL LDAVKSFRIRHEPGGLRLGHIPFOVAGV VGLKMPRVCLFGDTVMTASRMESNGEALKI HJSSETRAVLEEFGGFELLERGDVFWCKGK KVRTYWLLGERGSSTRG VGLKGERSSTRG QUTQVERSKALRKHAEEKFIEFDSGGCKK KVRTYWLLGERGSSTRG QUTQVERSKALRKHAEEKFIEFDSGCQEKK LUQTRVESLESGTRQDQEHQVELELLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGEHQUELLLGDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGEHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGEHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGEHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGEHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQDGHQVELLLRDDNE QUTQVERSKALRKHAEEKFIEFDSGCQEKK LHQUTGVESLESGTRQUSGCHQUTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG			<u> </u>	<u> </u>	sequence		
NGAFRVVLNYNGTSQELVANSGEKLNWPLG YPPPDPKGFONDEPACNQDHEISTEVLALV GSLSLLGILIVSFYYRKMQLEKELASELWAW WEDVEPSSLERHLRAGGSNTOSGNYGSI LITTEGQFQVFAKTAYYKGNLVAVKGNYRKR IELTRKVLFELKHMRDVQNEHLTRFVGACTD PPNICLITEVCPFGSLQDIENESTILDWIFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV DGRFVLKITDYGLESFEIDLPEGGHTVYAKK LWTAPELLRAMSPYYGSQAGDVYSFGILIGE ALRSGYFIVEGLLDSKEIERVTRGGPPR PSLALQSHLEELGLLMQRCWAEDPQERPPFQ QRLTLRKFNRENSNINDLLSRAMEQVANNL EELVEERTQAYLEEKRAEALLYQLDFHSVAE GLKRGETYQAEAFDSVITYFSDIVGFGPFR PSLALQSHLEELGLLMQRCWAEDPQERPFQ QRLTLRKFNRENSNINDLNLSRAMEQVANNL EELVEERTQAYLEEKRAEALLYQLDFHSVAE GLKRGETYQAEAFDSVITYFSDIVGFTALSAE STPMQVVTLINDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRIHACEVARMALAL LDAVRSFRIRHFPQEQLRLRIGHTGPVCAGV VGLKMPRVLGFGFTVATASRMESRGEALKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KYRTYWLLGERGSSTRG ASSTORMAN SERVSGLAGSIYREFRILWRYDEEFGGSAV MSERVSGLAGSIYREFRILWRYDEEFGGSAV MSERVSGLAGSIYREFRILWRYDEEFGGSAV GLTTPAYALGERGSSTRG ULTVAVLENLDSVFAODQEHQVELELLRDDWS QLTTPAYALGERGSSTRG ULTVAVLENLDSVFAODQEHQVELELLRDDWS QLTTPAYALGERGSSTRG ULTVAVLENLDSVFAODQEHQVELELLRDDWS QLTTPAYALGERGSFRWFGELSQFR LLVAVALENLDSVFAODQEHQVELELLRDDWS QLTTPAYALGERGFYKVTDAFNKSEISKHIEV QVAQETINNSTGSABAREKSEVQAIGHTPANSLGIFTP AGDGLTTPAQKGGTTGSSQWKFQELSQFR SHISLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEEBGFYKVTDAFNKSEISKHIEV QVAQETINNSTGSABREEKSEVQAIGESTFEL DMMCDLSGYKGSSTTTKGIGHNGAPPRNTESL FELSSAGSGLIGDVDGAGDLLGAGREVENLI LENTQLLETKNALNIVKNDLJAKVDELTCEK MARVLMENQVKERLMELGEAVRYTEMIR ASRENPAMQEKRSSIWGFFSSLPSSSSNTTK WARVLMENQVKERLMELGEAVRYTEMIR ASRENPAMQEKRSSIWGFFSSLPSSSSSNTTK VRAHVVAGDGAVALGFWSLPGKYKVTOMG QGENKMKNLPYTYRPLDEKOTSMKLWCA GGENKMKNLPYTYRPLDEKOTSMKLUCA GGENKMKNLPYTYRPLDEKOTSMKLUCA GGENKMKNLPYTYRPLDEKOTSMKLUCA GGENKMKNLPYTYRPLDEKOTSMKLUCA GGENKMKNLPYTYRPLDEKOTSMKLUCA SAEGVTOAATSISTNAASVYMBKPPEMEAB SEVDENNYTAEEATTATEONAGSAEDTVAGG GOTYTHATTDTGUQPPELSEVYQSSND SDAYKDQISVLPREQULVEEGAGKMSSLLPT MWGAQNGCLTVINSTCMATVYGNILDS FTVCNSIVLCTASVPGARETDYPAGEDLSEGG QUDKASTLPREGATATEONAGSAEDTVAGG SLESSLYWCTSTHASTRECHAVYPOYLOSND SDAYKDQISVLPREQULVEEGAGKMSSLLPT MWGAQNGCLTVIRSTCMATVYDENNOND SDAYKDQISVLADGULADG		ì	İ	l	ł		
YPPPDPKCGEDNEDÄCKODHLSTLEVLALV. GSLSLIGHLVSFIYEKMQLEKEL ASEL WRVW WEDVEPSSLERHLRSAGSRI TLSGRGSNYGSI LTTEGGFQVPAKLATY KGNLVAVKRVINERR IELTRKVLFELKHMRDVQNEHLTRFVGACTD PPNICILTEYCFRGSLQDILENESITLDWMFRY SLTNDIVKGMLELHNGAGSHGNLKSSNCVV DGRFVLKITDYGLESFRDILDFEQGHTVYAKK LWTAFELLRAASSPYRGSQAGDVSFGGILQE IALRSGVFHVEGLDLSFKEIJERVTRGEQPFPR PSLALQSHLEBIGLLMGRCWAEDPDSFFGILQE GRETQAFLEEKRAEALLYQDLFBYAE QUKRGETVQAEAFDSVTHYFSDIVGFTALSAE QUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVQAEAFDSVTHYFSDIVGFTALSAE GUKRGETVGAEAFDSVTHYFSDIVGFTALSAE GUKRGETVGAEAFDSVTHYFSLESSENTIS GUKRGETVGAEAFDSVTHYFSLESSENTIS GUKRGETTALSAEAFDETSCSGAEAFTALSAEAFD AGGGLTPDAQKGGFTAGSGWAKPGELSQFR SHTSLKDELSDVSQGGSKATTPASTANSDIVA TIPTDTPLKEENEGFVKVTDAPNISSISHIBV QVAQETRAVSTGSAENEEKSEVQAHESTPEL DMMKDLSGYKGSSTTRKGIENAEFDRINTESL FEELSSAGSGLIGDVDBGADLLGWGREVENLI LENTOLLETTMALINYKNDILAKVDBLTCEK DVLQGELBAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIFTAQRRCFTTRVE MAVLWARNQVKERLMELGEAV RVTEMIR ASRENAMGEKKRSSIVQFFSALPSSSSNITK KREPTVMALKNAMPTSHYKKRSSTLSQLP GOKSKAPDI-SEETEASLASRRCKREQVRQ QUKNAMPQKEDGRVQAFGWSLPQKVRQVTNG QGENKMKHLPVPYVLRPLDEKDTSMKLWGA VANINGEGKTPOGSGVVASVFYKVDYGNILDS FTVCNSHVLCTASVPGARETDYPAGEDLSESG QDVKASLCOSMTINSSAETDSLLIGGTVVGC SAEGVTGAATSPSTNGASPVORGNLDSTPVGSDLSESG QDVKASLCOSMTINSSAETDSLLIGGTVVGC SAEGVTGAATSPSTNGASPVORGNLEDKELNIKG SUNDNYTABENTAETATGNAGGABCDTVDIGS QUFKSHVLTNIFFSKKRRSSTLSQLP GOKSKAPDI-SEETEASLASRRCKREQVRQ QUFKSHVLTNIFFSKKRSSTLSQLP GOKSKAPDI-SEETEASLASRRCKREQVRQ GUKRKINLTYVPYLRIPLEKDTVYNCFOILDS FTVCNSHVLCTASVPGARETDYPAGEDLSESG QDVKASLCOSMTINSSAETDSLLIGGTVVGC SAEGVTGAATSPSTNGASPVORKPCBLERISKAD SEVDENDYTAEBEATEATGNAGGAEDTVDIS GT		ŀ					
GSI.SLI.GILIVSFITYRIMQLEKELASEL WY WEDVEPSSI.ERHIRSAGSSILT.SGRGSNYGSI LTTEGOFOVFAKTAYYKGNI.VAVKRVINKER IELTRKVI.FELKHMIPOVQNEHI.TREVGACTD PPMICILTEYCERGSI.QDILENESITLDWENGY SLTNDIVKGMLFLHINGAICSHIGOTAKSNCVV DGRFVLKITDYGLESFRILDPEQGHTVYAKK LWTAPELLRMASPYVRGSQAGDVYSGILIG ELALRSGVHFWEGLDSFKEILERVITRGEQPPR PSI.ALQSHLEELGLLMQRCWAEDPQERPFOQ QIRLTLRKFYRERSSNILDNILSRMEQYANNI. EELVEERTQAYLEERKKABALLYQILPHSVAB QIKRGETVQAEAPDSVITYISDIGTALSAE STFMQVVTLIADLYTCFDAVIDNEDVYKVCY IGDAYMVVSGLYVRORGHIACEVARMALAL LDAVRSFRRHRRQEQLRIRGHTGPVCAG VGLKMPRYCLFGDTVNTASRMESNGEALKV HLSSUETKAVILEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG VGLKMPRYCLFGDTVNTASRMESNGEALKV HLSSUETKAVILEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG APPGRPVRLFTMELEDGVYQEEFGGSGC VGLKMPRYCLFGESSTRG DAPGRPVRLFTMELEDGVYQEEFGGSGC KVRTYWLLGERGSTRG VGLKMPYCLFGESTRG LVVAVLENLDSVFAQDQEHQVELELLRDDV QLTGYPEREALRKHAEEKEFFEDSQEQEKK DLQTRYBELESQTTQLELKAKNYADQISLEE REAELKENNALHGRHTEMHENYMEHLEET REAELKENNALHGRHTEMHENYMEHLEET REAELKENNALHGRHTEMHENYMEHLEET REAELKENNALHGRHTEMHENYMEHLEET REAELKENNALHGRHTEMHENYMEHLEET REAELKENNALHGRHTEMHENYMEHLEET DMKDLSGOQLESTAHSRIRKEEPSIGIPPP AGDGLITPDAQKGGETPGSEQWKFQELGOP SHTSILDELSDVSQGGSKATTFASTANSDVA TIPTTDTPLKEENGGFVKVTDAFNKSEISKHEIP QVAQETRANTSGASAENEKSEVQAIESTFEL DMKDLSGVKGSSTFYKGIENKAADRINGEN TIPTTDTPLKEENGGFVKVTDAFNKSEISKHEIP QVAQETRANTSGASAENEKSEVQAIESTFEL DMKDLSGVKGSSTFYKGIENKAADRINGEN ARREWPAMQEKRSSIWGFFSKLFSSSSTTITK KPEPPYNILKYNAPISHYTPSVKRKEEEELKAA RAEEDARQKAKDDDDSDIPTAQRKRFTTRVE MARVLMERNGYVERIMELGAGKVKGVVTNG QGENKMKNILPVPVYLRFLDEKDTSMKLWCA VGVNLSGCKTRGGGSVVGASVTVENMIR ASREWPAMQEKRSSIWGFFSKLFSSSSSTITK KPEPPYNILKYNAPISHYTPSVKRKESTI.SQLP GDKSKAFPFISHSTNSAFTDSALCHOOR VKAHVQKEDGFVKYROYTNG QGENKMKNILPVPVYLRFLDEKDTSMKLWCA VGVNLSGCKTRGGGSVVGASVTYCHOM VKAHVQKEDGFVKYROYTNG QGENKMKNILPVPVYLRFLDEKDTSMKLWGA SEVDENVFTAEEATRALTENAGAGKSAGDTUNG SEVDENVFTAEEATRALTENAGAGKSSILOKLDOGLKGQCELKNIL GELSLVWICTSTRATKAVLULDAGGTVVGCC SAEGVYGAATSPSTNGASPOMRPPEMEARN SEVDENVFTAEEATRALTEGNAGASAEDTUNG SAYLDDIGSVLPREGQLKYROYTHORLUNG VKAHVQKEDGFVYVSTENGARCHISIKCM SILSTHYKGGKTLDSTLATHAT							
WEDVERSLEHLRSAGSÄLTLSGRGSNYGSL LTTEGGGOVPAKTAYYKGNU KRYNNERR IELTRKULFELKHMRDVQNEHLTREVGACH PPINICILTEVCRESCJ, OELENEESITLD WIMFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV DGRPVLKITDYGLESFRDLDFEGGHTVYAK LWTAPPELRMASPVRGSQAGDVSFGIILQE IALRSGVFHVEGLDLSFKEHERVTRGGEPFPF PSLALQSHLEELGLLMQRCWABDPQERPPFQ QIRLTLRKFNRENSSNILDNLLSRMEQYANDL EELVEERTQAVLEEKRKABALLYQILPHSVAE GUKRGETVQAEAFDSVTIYFSDIVGFTALSS STFMQVVTLLNDLTYCTPAUIDNFDVYKVET IGDAYMVVSGLPVENGRIHACEVARNALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGFALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGFALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGFALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGFALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGFALL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNSHENGER QLTQVTERFGALRHGFGFELLERGVYVQEERCGSGAV MSERVSGLAGSTVREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQERQVELELLEDDVENKGKG KVRTYWLLGERGSSTRG DAPGRPPVELFFGGFELLERGVYVQEERCGSGAV MSERVSGLAGSTVREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQERQVELELLEDDVA QLTRVESLESQTRQLELKAKNYADQISLEE RAELKKEYNALHGRHTEMINFYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFTPL AGGGLTTPDAVGGGGTTGSGEVWFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANTSVI TIPTDTPLKEENGFGVAVTDATNYSTGSAENEERSEVQAHESTFEL DMDKDLSGVKGSSTPTKGEIRKAPRTESLE FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLETINNALNIVANDLAKPRINTESL FEELSSAGSSLIGDVDEGADLLGMGREVENLI LENTQLEFTINNALNIVANDLAKAR RAEAEDARQKAKDDDDSDIPTAQRKKESKINIEV QVAQETRNSVTSGSAPTNEERSENFRINSLEFELRX RAEAEDARQKAKDDDDSDIPTAQRKKFTRVE MARVLMERRQYKERLMINIVANDLAKAR AGREDPANQEKKRSSIWQFFSRLFSSSSTITTL FEELSSAGSSLIDKLDDGLKFGPKKFTVVE MARVLMERRQYKERLMINIVANDLAKAR AGREDPANGEKKRSSIWQFFSRLFSSSSTITTL FEELSSAGSSLDKLDDGLKFGQKELKNI EELSLVWICTSTTASTKVTLTUDKVVOGG VKAHVQKEDGGVVAGSVYGNSVYDVAGL TGGSKQRSASQSSLDKLDQELKFQQKELKINI SENDENVPTAEBATTATEGNAGSAEDTVADIS GYDYSTEHVYTDEVTSARKCLHSIKKUT SDAYKDQISUNDENGLJKREAGKMSSILDY MWLGAGNGCLYVHSSVAQWRGCKSRDARDSSDAYCDG DLSNYNHLDLGGRHTHSIRCMTVHDKVWGC VRNKIYVVQRKAMKIEKSPDAHPRESGYGN LAWVYDRAM		1		l			I
LTTEGGFQVFAKTAYYKGNI.VANKRINRISE BELTRKULFELKHMRDVQNEHLTREVGACTD PPNICIL TEVCRGSL.QDILENESITL.DWMFRY SLTNIDVKGMLFLHNGAICSHOTLKSNCVV DGRPVLKITDYGLESFRDLDPEQGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGILGE BALRSGVFHVEGLDLSFKEILERVTRGEQPFPR PSLALQSHLEELGLLMQRCWAEDPQERPFY QIRLTLRKFYRERSSNILDNILLSRMEQYANNL EELVEERTQAYLEEKKABALLYQILPHSVAE QLKRGETVQABAFDSVTIYFSDIGFTALSAE STFMQVYTLLNDLYTC-CDAVIDINFDVYKLY BELVEERTQAYLEEKKABALLYQILPHSVAE QLKRGETVQABAFDSVTIYFSDIGFTALSAE STFMQVYTLLNDLYTC-CDAVIDINFDVYKLY GOLDAWNYSGLYBROENLARGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSETKAVLEFGGFELELRGDVEMKOKG KVRTYWLLGERGSSTRG ASSENSGLAGSIYREFERLIVRYDEEVVERL LUVAVLENLDSVAPQQDEHQVELELLRDDINE QLTQYEREKALRKHAEKFIFEFEDSQEGEKK DLQTTQYEREKALRKHAEKFIFEFEDSQEGEKK DLQTTQYEREKALRKHAEKFIFEFEDSQEGEKK DLQTTYSELSGYRQLELKARNYADQISILEE RAELLKESYNALHORHTEMINNYMEHLERT KLHQLSGSDQLESTAHSRINKERIGIFPLP AGGGLLTPDAQKGGETFGSEQWKFOELSOS SHISTLXDELSDVSGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDATYNKSEISKHIEV QVAQETRINNSTGSAENEKSEVQAIESTEL DMDKDLSGYKGSSTTYRGERKARDERELEL DMDKDLSGYKGSSTTYRGERKARDERELEL DMDKDLSGYKGSSTTYRGERKARDERELEL DMDKDLSGYKGSSTTYRGERKARDERELEL FEELSSAGSGLIGDVDEGADLLGMGREVENIL LENTQLETKNALNIVKNDLLAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEELRAF RAEABARQKAKDDDSDIPTAYRKESISNIFTK KPEPPVILKYNAPTSHVTPSVKKRSSTLSQLP GDCSKAFPTISEETEALSARSREQKEQVYQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVITVTSVKKRSSTLSQLP GGCSKAFPTISESTETASLASREDKEDSTSNKLWCA VGNILSGKKTRGGSSVYGASVFVDNGL GGENKMKNLPVITVTLRIDERDTSNKLWCA ASREPNAMGEKRSSIWGFFSKSSSNITTK KPEPPVILKYNAPTSHVTPSVKRSSTLSQLP GGCSKAFPTISESTETASLASREDKEDEVQN VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVITVTLRIDERDTSNKLWCA SAEGVTGAATSPSTNCASSYMDKPPEMABAN SEVDENVITAEBATGATTAGTACKGARDTVADIS QCOVYKSALCGSMTNSSAETTBLLGGITVVGC SAEGVTGAATSPSTNCASSYMDKPPEMABAN SEVDENVITAEBATGATTAGTACKGARSSLITP MWLGAQNGCLYVHSSVAQWRKCLISIKLKD SILSIVHYKGIVLVALADGTLAFHRGVOYGC YNNKIYVVQEKAMKLEKSFDAHPRKESGYG LAWYGOVWWSILDSTLRLYHAHTYQHLQ LAWVGOVWWSIRLDSTLRLYHAHTYQHLQ							
IELTRKVLFELKHMRDVQNEHLTRFVGACSI PPMICIL TEVERGES, QOILENESITLOWMFRY SI.TNDIVKGMLFLHNGAICSHGNLKSSNCVL DGRPVLKITDYGLESFENLDPEQCHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGIILQE IALRSGVFHVEGLDLSPKEIIERVTRGGPPFPF PSI.AL, GSHLEELGLLMQRCW.ABDPQERPFPF QIRTILRKFRRENSSNILDNLLSRMEQVANIALL ELLVERTQAYLEEKRKABALLYQILPHSVAL GLKRGETVQAEAPDSVTYIYSDIVGFTALSA- STFMQVVILLNDLTYCFDA.VIDNFDVYKVET IGDAYMVVSGLPVENGRLHACEVARMAL- LLDAVRSFRRHRPGSCJLRICGHITGPVCAGL VGLKMPRYCLFGDTVNTASRMESNGGALIKI HLSSETKAYLLEFGGFELLERGHEMGFALKI HLSSETKAYLLEFGGFELLERGHEMKOKG KVRTYWLLGERGSSTRG DAPGPPVRLFHBLEFGGFELLERGHEMKOKG KVRTYWLLGERGSSTRG DAPGPPVRLFHBLEFGGFELLERGHEMKOKG KVRTYWLLGERGSSTRG DAPGPPVRLFHBLEFGGFELLERGHEMKOKG KVRTYWLLGERGSSTRG LVVAVLENLDSVFAQDQEHQVELELLRDDW GLTQYEREEAURKHAEEKFEFEDSGEGEKK DLQTRVESLESQTRQLELKAKNYADQISLE ERAELKKEYNALHGRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPL AGGGLTTPAGNGGGTTGSGEWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSUVA TIPTDTPLKEENGFVKYTDAPKSEISKHIEV QVAQETRNYSTGSAENEEKSEVQAHESTFEL DMDKDLSGVKGSSTPTKGIENKAPRTYSEL DMDKDLSGVKGSSTPTKGIENKAPRTYMEH DMDKDLSGVKGSSTPTKGIENKAPRTYMEH DMDK							
PPNICILTEYCPRGSLQDILENESITLUMMERY SLINDIVKGMILFLINGALCSHGNLKSSNCVV DGRPVLKITDYGLESFRDILDPEGGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGILLOG IALRSGVFHYEGLDLSPKEIIERVTRGEQPPPR PSLALQSHLEELGLIMQRCWAEDPQERPPPQ QIRLTLRKFINENSSNILDIALISMEQYANNI. EELVEERTQAYLEEKRKAEALLYQILPISVAE QLKRGETVQAEARDSVTITYSDIVOFTALSAE STPMQVVTLLNDLYTCFDAVIDNIPDVYKVET IGDAYMVVSGLPVNRGRLHACEVARMALAL LDAVRSFRIRHPQSGLRLRGIHTGPVCAGV VGLKMPRVCLFGGTVNTASRMESNGEALKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVVQEEPGGSGAV MSERVSGLAGSTYREFERLINKYDEEVVKLIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTIQVEREKALRKHAEEKFIEFEDSOGEKK DLQTRVSLESGSTROLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIENYMEEVVKLIP KLHQLGSGSDLOSTAHSRIKKERPISLGIFPLP AGDGL1TPDAQKGGETPGSEQWKPQELSQPR SHTSLXDELSDVSQGGSKATHSTANDSVA TIPIDTPLKEENBGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIESTPEL DMDKDLSSVKASSTPTKGIENKAPDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLETKNALNIVKNDLJAKVDELTCEK DVLQGELAVKQAKLKLEEKNRELEEELRA RAEADARQKAKDDDDSDIPAGKRSTIVGFSRLFSSSNTTK KPEPPVILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAPDELSEETEASLASRQKREQYNQ VKAHVQKEDGRVQAFGWSLPQKVKQVTNG QGENKMKNILPVPVYLRPLDEKDTSMKLWCA VGVALSGGKTRDGGSVVGASVFYKDVAGLD FTVCNSHVLCIASVPGARETDYPAGEDLESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC QSRKMKNILPVTYPTPLERGDTSMKLWCA VGVALSGGKTRDGGSVVGASVFYKDVASND SDAYKDDISVLFREATATATEDNAGAEDTVADS FTVCNSHVLCIASVPGARETDYPAGEDLESG QVDKASLCGSMTSNSSAETDSLLGGITVGC QSRGWTGATSTSTNGASPWMDKPPEMAEN SEVDENVPTAEAGATTATAPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHVKGTVLVALADGTLAIPHRGYDGQW JSLSTYHLGLIGGPHTSISKGCHTVVHDKUVCG YRNKIYVVQPKAMKERSFDAHPRKESQVRQ LSNYHLLGLAGPHTSISKGRATLVHAHTYQHDGU GULSDYHILLDLAGPHTSISKGRATLVHAHTYQHDGU GLANGGGRUPSISLDSGTLICLTHAITTONTAL	1	1					1
SLINDIVKGMLFLIHIĞAICSHGRLIKSNCVV DGRPVLKITIVQLESPRDIDPEQGHTVYAKK LWTAPELLRMASPPYVRGSQAGDVYSFGIILQE IALRSGVFHYEGLDLSPKEIIERVTRGEQPPR PSLALQSHLEELGLLMQRCWAEDPQERPPG QIRLTLRKPNENSSNILDNLISRMEQYANNL EELVEERTQAYLEEKRKACHVQILPHSVAE QLKRGETVQAEAFDSVTIYPSDIVGFTALSAE STPMQVVTLINDLYTCFDA VIDNEDVYKVET IGDAYMVVSGLPVRNGRLHACEVARMALAL LDAWSSPRIRHRPOEQURLRIGHTGPVCAGV VGLKMPRYCLPGDTVNTASRMESNGEALIKI HLSSBETKAYLEEFGGFELELRGDVEMKGKG KNTTYWLLGERGSSTRG QLTQVEREKALRHAEEKEKFEEDSGQEKK DLQTRVESLESQTRQLELKAKNYADQISLEE REABLKEYSALAGSTYREFFELIVRYDEEVVKELIF LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTQVERSLAKRKHAEEKFEEDSGQEKK DLQTRVESLESQTRQLELKAKNYADQISLEE REABLKEYSHALHQRITEMHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGGFPLS SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDANNKSEISKHIEV QVAQETRNVSTGSAENEEKSQAIESTFEL DMDKDLSGYKGSSTPTKGIENKAPDRITEN FEELSSAGSGLIGVDEGADKTESLE FEELSSAGSGLIGVDEGADKGREVENIL LENTQLLETKNALNIVKNDLLAKVDELTCEK DVLQGELEAVKQAKLLEEKERLEELDRAK RAEAEDARQKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEELBLRAK RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERSLEBELBGGTVVQG GENKMKHLPPVYYLEPLDEKDTSNKLWCA GYGNESSCAFTSTSTSTOGAFYMDKPPSPRAEAD SEVDENSTAEBERAGKREDYDGNU DSNYHLLDLGRPHTSTECHTAGASAEDTVAGS DVKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGAATSPSTNGASPAMDKPSPEMAEN SEVDENVPTAEBARTATTEGNAGSAEDTVADIS DAYTHULDLAGPHISTRCMTVVHDKVWG ULSNYHLUDLAGRPHISTRCMTVVHDKVWG ULSNYHLUDLAGRPHISTRCMTVVHDKVWG ULSNYHLUDLAGRPHISTRCMTVVHDKVWG ULSNYHLUDLAGRPHISTRCHTVVHDKVWG ULSNYHLUDLAGREAAAADTLAJHTRYDOQW DLSNYHLLDLAGRPHISTR)]]			
BORFVLKITDYGLESFRDILDFEGGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGILLOG IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR PSI.ALQSHLEELGLLMQRCWAEDPQERPFPQ QIRLTLRKFNRENSSNILDNLSRMEQYANNL EELVEERTQAYLEEKRKAEALLYQILPISVAE QLKRGETVQAEAFDSVITYSDIVGFTALSAE STFMQVVTLLNDLYTCFDAVIDNEDVYKVET IGDAYMVVSGLPVNRGHI HACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGHTGPVCAGV VGLKMPRYCLPGDTVNTASRMESNGEALKI HLSSETKAVLVEEFGGFELELRGDVEMKOKG KVRTYWLLGERGSSTRG BOPGRPPVELTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYREFRRIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQKK DLGTRYESLESQTRQLELKAKNTADGISLEE REABLKKETVAHALHQRITEMHINYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIPPLP AGDGLTTPDAYGAGGTTPGSGWKFQELSQFR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPITDTPLKEENEGFVKVTDAKSEISKHIEV QVAQETRAVSTGSAENEKSEVQAIESTPEL DMDKDLSGYKGSTPTKGIENKAPDRINTSL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALINVKNDLJAKVDELTCEK DVLQGELEAVKQAKKLEEKKREELEELKRA RABAEDARQKAKDDDDSDIPTAQKRKTRTVE MARVLMERNGYKERLMBLGAAVR WTEMIR ASKENPAMQEKKRSSIWQFFSRLFSSSSNTIK KPEPPVNLKYNAPTSHVYTPSVKKRSSTLSQLP GDGSKAFDELSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYVLRPLDEKDTSMKLWCA VGVALSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQEAKQVTNG QGENKMKNLPVPYVLRPLDEKDTSMKLWCA VGVALSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQOKELKNQ VGVASLCOSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAN SEVDENVPTAERAGASAEDTVYSSND SDAYKDOISVLPREQAVKSSLLPLT WWLGAGAOCCLYVHSSVAQWRCCLHSIKLKD SILSIVAVKGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLAGPHISIRCRMTVVHDKVWGL SILSIVAVKGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLAGPHISIRCRMTVVHDKVWGL SILSIVAVKGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLAGPHISIRCRMTVVHDKVWGL AWVGDGVWYSIRLDSTLTLALTHANTYQHLQ DLSNYHLLDLAGPHISIRCRMTVVHDKVWGL LAWWGDGVWYSIRLDSTLTLALTHANTYQHQLQ ULSNYHLLDLAGPHISIRCRMTVVHDKVWGL LAWWGDGVWYSIRLDSTLTLALTHANTYQHCL AWWGDGVWYSIRLDSTLTLALTHANTYQHCL AWWGDGVWYSIRLDSTLTLALTHANTYQHCL AWWGDGVWYSIRLDSTLTLALTHANTYQHCL			ļ		İ		L
LWTAPELLRMASPPYRGSQAGDIVYSFGIILQE IAIRSCVPHYCGLDLSPKEIDGLOPPRP PSIALQSHLEELGLLMQRCWAEDPOERPPPQ QIRLTLRKPINERNSNILDNILSRMEQYANNL EELVERTQAYLEEKRAEALLYQILPHSVAE QLKRGETVQAEAFDSVTIYYSDIVGFTALSAE STPMQVVTLINDLYTCFDAYIDNFDVYKVET IGDAYMVVSGLPVRNORLHACEVARMALAL LDAYRSFRIRHPQGQLRIKGIHTGPVCAGV VGLKMPRYCLPGDTVNTASRMESNGEALKI LLSSETKAVLVEEFGGFELBLRGDVYEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLFTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFREILVRYDGEVYKELIP LVVAVLENLDSVFAQDQEHQVELELLADDNE QLITQYEREKALRKHAEEKFEFEDSQECKK DLQTRVESLESGTRQLELKAKNYADQISLEE REABLKEYNALHGRITEMNYMEHLERT KLHOLSGSDOLESTAHSRIKEERPISLGIFPLP AGDGLLTPDAVKGGETPGSEQWKFQLLSQPR SHTSIKDBLSDVSQGGSKATTASTANSDVA TIPTDTPLKEENEGFVKVTDATNISESISHIEV QVAQETRNVSTGGSAENEKSEVQAIIESTFEL DMDKDLSGYKGSTPTKGIENKAFDRISSHIEV QVAQETRNVSTGSAENEKSENQAIIESTFEL DMDKDLSGYKGSTPTKGIENKAFDRISSHIEV QVAQETRNVSTGSAENEKSENQAIIESTFEL LENTQLLETKNALNIVKNDILAKVDELTCEK DVLQGELEAVKQAKKLEEKRRELEELRSTLSQLP GUKSKAFDFLSSETEASLASRREQKREQVRQ WKAHVQKEDGRVQAFGWSLPQKVKQVTNG QGENKMKHLPVYYLEPJDEEDTSMKLWCA VGVALSGGKTRDGGSVVGASVFXKDVTNG QGENKMKHLPVYYLEPJDEDTSMKLWCA VGVALSGGKTRDGGSVVGASVFXKDVTNG QGENKMKHLPVYYLEPJDEDTSMKLWCA VGVALSGGKTRDGGSVVGASVFXKDVTNG QGENKMKHLPVYYLEPJDELDSTDSMKLWCA VGVALSGGKTRDGGSVVGASVFXKDVTNG QGENKMKHLPVYYLEPJDELDSTDSMKLWCA VGVALSGGKTRDGGSVVGASVFXKDVAGLD TEGSKORSASSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFONILDS FTYCNSHYLCLASVFQARETDYPAGEDLESGG QVVKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPMDKPPEMAEN SEVDENVPTAERAGATSATTDVPAGEDLESGG QVVKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPAMDKPPEMAEN SEVDERAVFTAEGARTSSVANGASPATVONSND SDAYKDDISVYPRIDDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRCCLHSIKLKD SILSVHVKGIVLVALADGTLAIPHRGVORQW DLSNYHLLDLGRPHHSIRCKHTVVHDKVWGN LAWVGDGVWVSIRLDSTLILLTHANTTOVHDKWCG YRNKIYVVQFKAMKERSFDAHPRKESQVRQ LAWWGDGVWVSIRLDSTLILLTHANTTOVHDCWWCG YRNKIYVVQFKAMKERSFDAHPRKESQVRQ LAWWGDGWVVSIRLDSTLILLTHANTTOVHDC			l				1 .
I ALRSGYFHVEGILDLSPKEIJREVTRGEOPPFF PSLALQSH, BEG, GLLMQRCWAEDPGEPPFQ QIRLTLRKFNRENSSNILDNLLSRMEQYANNL EELVEERTQAYLEEKKRAEALLYQILPHSVAE QLKRGETVQAEAPDSVITYSDINGFTALSAE STPMQVVTLLNDLYTCFDAVIDNFDYKVET IGDAYMVVSGLPVRNGRHACEVARMALAL LDAVRSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYGLFGDTVNTGSPENGFALKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYRFFERLIVRYDEVVKELH LVAVALE.HLDSVFAQDGEHQVELELLRDDNE QLTTQYEREAJRKHAEEKFIEFGDSQEKK DLQTRVELSGSTRG GLENGHLENDNE QLTTQYEREAJRKHAEEKFIEFGDSQEKK DLQTRVELSGSTRG HEAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERFISLGIFFLP AGGGLTFDAQKGGETFGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENGGFVXVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEKSEVQAIBSTFEL DMDKDLSGVKGSSTRT KGENKAPDRNTESL FEELSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKALAINVKNDLIAKVDELTCEK DVLQGELEAVRQAKLKLEEKNRELEEELBKA RAEADARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERGVYKERJMELGGAVWETMIN ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVBLKVNAPTSHVKMLGLGAVWETMIN WARVLMERGVYKERJMELGGAVWETMIN GGENKMKNLEVAVARTSHVKJVADTSHVKJVADSHSTVKJVADGGILGD TEGSKQRSAGSSLDKLEDGTSTRANGV VKAHVQKEDGRVQAFGWSLPQKYKQVNG QGENKMKNLEVATATERGMAGSAEDTVADIS GGENKMKNLUCLASVPGARETDVPAGEBLSSES QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPPMEAEN SEVDENVPTAEEATEAGMAGSAEDTVADIS GJGTYTHENTTDPLGVQPEDLSSG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPPMEAEN SEVDENVPTAEEATEA-TEGMAGSAEDTVADIS GJGTYTHENTTDPLGVQPEDLSSG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPPMEAEN SEVDENVPTAEEATEA-TEGMAGSAEDTVADIS GJGTYTHENTTDPLGVQPEDLSSG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPPMEAEN SEVDENVPTAEEATEA-TEGMAGSAEDTVADIS GJGTYTHENTTDPLGVQPEDLSSG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPPMEAEN SEVDENVPTAEEATEA-TEGMAGSAEDTVADIS GJGTYTHENTTDPLGVQPEDLSSG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAGGVTGAATSPSTNOASSPVMDKPCLAISIKLKD SILSVHVKGVTLALADGVATLAIFHGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCC YRNKITYVQRAMKIEKSSPDAHPRESGOVRO LAWVGBQVWVSIRLDSTRLATHATYOHLQ			ļ	!			1
PSIALQSHLEELGILMQRCWAEDPQERPPPQ QIRLTLRKFNRENSSNILDNLLSRMEQYANNL EELVEERTQAYLEEKKAEALLYQILPHSVAE QLKRGETVQAEAFDSVTIYFSDIVGFDLYSSVAE QLKRGETVQAEAFDSVTIYFSDIVGFDLYSVAE GLKRGETVQAEAFDSVTIYFSDIVGFDLYSVAT IGDAYMVVSGLPVRNGRIHACEVARMALAL LDAVRSFRIRHRPGEQLIKGIHTGPVCAGV VCLKMPRYCLFGDTVNTASRMESNGEALIKI HLSSETKAVLEFGGFELELRGDVEMKGKG VCLKMPRYCLFGDTVNTASRMESNGEALIKI HLSSETKAVLEFGGFELELRGDVEMKGKG KYRTYWLLGERGSSTRG SERREILVRYDEEVYKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIFSDQEQEKK DLQTRVESLESQTRQLELKAKNYADGISILEE REAELKKEYNALHORHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSKEPISLGIFPLP AGDGLLTPDAQKGGETTGSEQWKFQELSQFR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTIGAENEKSEVQAIESTFEL DMDKDLSGYKGSSTPTKGIENKAFPRNTESL FELSSAGSGLIGDVDEGADLLGMGEVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWGFFSALFSSSNTTIK KPEPPVNILKVNAPTISHVTSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASREQKEQYRQ VKAHVQKEGGRVVAGAVSVFXUNGD GGENKMKNLPVPYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLI EGSKQRSASSSLNELDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFGNILDS FTVCNSHVLCASVFGARSTDYPAGEDLSEG QVDKASLCGSMTSNSSAETDSLLGGITVVGG SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENNYTAEEVATEATEONAGSAEDTVUDIS QTGYYTEHVTTDPLGVQUEPDLSPGYQSSND SDAYKDQUSVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWEKCLHSIKLKD SILSIVHVKGIVALAJGRICHSIKLSD SILSIVHVKGIVALAJGRICHSIKLSD SILSIVHVKGIVAVALAGRICHSIKLSD SILSIVHVKGIVAVALAGRILAFRRENGOVRO DLSNYHLLDLGRPHISIRCMTVVHDKVWCG YRNKIYVVQRKAMKIEKSSPDAJFPRESOVRO LAWVGBQWVSIRLDSTLRJYHAHTYQHLQ		1		•			,
ORLITEKENPRENSSNILDNILSERMEQYANNIL EELVEERTQAYLEEKRKAEALLYQIIPHSVAE QLKRGETVQAEAFPSVTIYFSDIVGFTALSAE STPMQVVILLNDLYTCPDAVIDDRPDYKVET IGDAYMVYSGLPVRNGRIHACEVARMALAL LDAVMSFRIRHRPGEQLRLRIGHTGPVCAGV VGLKMPRYQLFQDTVNYSGMEANESNGEALIKI HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG BAPGRPPVILTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLINRYDEEVVKELP LVVAVLENLDSVFAQDEVOELELLRDDNE QLITQYEREKALRKHAEEKPIEFFDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHEMINYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIPPLP AGGGLITPDAQKGGETSQWKPGELSQPK SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNSEISKHIEV QVAQETRANVSTGSAENEEKSEVQAIESTFEL DMDKDLSGYKGSSTPTKGIENKAPRENTESL EPELSSAGSGLIGDVDEGADLLGMGREVENIL LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEELRKA RAEAEDARGKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERLMELGEAVRYEMI ASRENPAMGEKKSSIVGFSRLPSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKATDHI-SEETBASLASREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYVLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVPYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFGNILDS FTVCNSHVLCLASVPGARETDYPAGEDLSEG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGTTVVGC SAEGVTGGAATSPSTNGACKLISIKLKU SILSIVHVKGRÜVALALAGGTALFHRGVDGOW DLSNYHLLDLGRPHSIRCMTVVHDKVWCG YRNKTYVVGRYAMKIEKSFDAJFPRESQVRG LAWVOEDGWVSIRLDSTRLHYHAHTYQHLQ LAWVOEDGWVSIRLDSTRLHYHAHTYQHLQ LAW	•	ļ	}				
EELVEERTQAYLEEKRKAEALL YQILPISVAE QLKRGETVQAEAFDSVTIYFDIVGFTALSAE STPMQVVTLLNDLYTCPDAVIDNFDYYKVET IGDAYMVVSGLPVENGRIHACEVARMALAL LDAVRSFRIRHRPQEGLERIGHTOPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALIKI HLSSETKAULEFGGFELERGDVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAFGRPFVRLFTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYRFERERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYERBEARKHAEKERFIEFDSQEÇEKK DLQTTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMHINYMEHLERT KLHQLSGSDQLESTAHSRIKKERFISLGIPLP AGDGLLTPDAQKGGFTGSSQWKKPGELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNSEISKHIEV QVAQETRANVSTGSAENEKSEVQALIESTPEL DMDKDLSGYKGSSTPTKGIENKAFPRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEELKKA RAEAEDARQKAKDDDDSTPTAGRKRTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASBENPAMQEKKRSSIWQFFSALFSSSNTTK KPEPPVNILKVNAPTISHVTSVKKRSSTLSQLP GDKSKAPFFSSEETEASLASRREQKREQYKOQ VKAHVOKEGGRVQAFGKEGYPRQ GGENKMKNLPVPVYLRPLDEEDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ GGENKMKNLPVPVYLRPLDEEDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASSSLDKLOQEKEQQKELKNQ GELSSLVWICTSTHSATKVLIIDAVQFGNILDS FTVCNSHVLCLASVFGARETDYPAGEDLSESG QVDKASLGGSMTSNSSAETDSLLGGITVVGG SAEGVTGAATSPSTNGASPWDKYPEMEABN SEVDENVPTAEEVATEATEGNAGSAEDTVVDIS GJGYYTEHVPTDPLGVQUEPELSPYQSSND SDAYKDQUSVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWBKCLHSIKLKD SILSIVHVKGRIVALALAGGTLAIFRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKITVVQRKAMKIEKSSPDAJFPRESOONG LAWVGBOWVSIRLDSTRLYHAHTTYCHLQ							
QLKRGETVÖAAAFDSVTIYYSDIVĞFTALSAE STPMQVVTILNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRIHACEVARMALAL LDAVRSFRIRHPQEQLRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSETKAVLEFEGGFELELRGDVEMKGKG KVRTYWLLGERGGFELELRGDVEMKGKG KVRTYWLLGERGSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYREFERLIVRYDEEVVKELIP LVVAVLVENLDSVFAQDEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFDSOGQEKK DLQTRVESLESQTRQLELKAKNYAQQISILEE RRAELKKEVNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIPPLP AGGGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPIDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAHESTPEL DMMKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWOFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYYLRFLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASOSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLGGITVVGC SAEGVTGAATSPSTNGASPVMDKYPEMAEN SEVDENVPTAEEATEATEGNAGSAEDTVODIS QTGYYTEHYPTDPLGVVQPEDLFPVQSSND SDAYKDQISVLPNEQDLVREEAQKKNSSLPT MWLGAQNGCLYVHSSVAQWRCLHSKIKL SILSVHVKGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAAHPRKESQYRQ VLAWVGDCVWVSIKLUSTIRLYHAHTYQHLQ							
STPMQVVTILNDLYTCFDAVIONFDVYKVET IGDAYMVVSGLPVRNGRLHACEVARMALAL LDAVRSFRIRHPQEQELRLRIGHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALIKII HLSSETKAVLEEFGEFELRGDVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYREFRELIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQVEREKALRKHAEKEFIEFFEDSGEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE RBAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISIGIPFLP AGGGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNYSTGSAENEEKSEVQAIIESTTEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELAVKQAKLKLEEKNRELEEELEKAA RAEAEDARQKAKDDDDSDIPTAQRKRTTRVE MASVLMERNQVKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWGFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHTYPSVKRSSTLSQUP GDKSKAFDFLSSETEASLASRREGKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSL VWICTSTHSAFTKVLIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLGGITVVGC SAEGVTJGAATSSTSTINGASPVMDKYPEMAEN SEVDENVPTABEATEATEGNAGSAEDTVADIS GJGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKKSSLLPT MWLGAQNGCLYVHSSVAQWRCLHSKIKL SISTIMVRGGIVLVALADGTLAIPHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFTALFHAHTYQHLQ LAWVGDGVWVSIKLLDSTLRLYHAHTYQHLQ		1		1			•
GIDATMYVSGLPVRNGRLHACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV							
LDAVRSPRIRHRPQEQLRLRIGIHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEALKI HLSSUETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG DAPGRPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDINE QLTTGYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEKSEVQAILESTFEL DMMKDLSGYKGSSTFTKGIENKAFDRTTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALINVKNDLLAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEELERKA RAEAEDARQKAKDDDDSDIPTAQRKFTRVE MARVLMERNQYKFRLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNILKYNAPTISHVTPSVKKRSSTLSQLP GICKSKAAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QEENKMKNLPYNIPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ ELSSLVWICTSTHSATKVLIIDAVQFGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSEG QVDKASLGGMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATTEATEGNAGSAEDTVDIS QTGVYTHVPTPDEJQVPGDELSPQSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SLSIVIVKGGVLVALADGTLAIFRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQFKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWSRLDSTLRLYHAHTYQHLQ			l				
VGLKMPRYCLFGDTVNTASRMESNGEALIKI HLSSETKAVLEEFGGFELELRGDVEMKKGK KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSTYREPELIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLTTQYEREKALRKHAEEKFEFEDSQEQEKK DLQTRYVESLESQTYRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKERPISLGIPPLP AGDGLLTTPDAQKGGETPGSEQWKYQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIBNKAFDRNTTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLLAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAFAEDARQKAKDDDSDIPTAQRKFFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIVQFFSLFSSSSNTTIK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVQASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQFGINLDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGIAGSAEDTVDIS GTVYTEHVFTIPDLGNVGPBLESPYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SLSIJVHVGGTULVALADGTLAIFRGVDGQW DLSNYHLLDLGRPHHSIRCMTVYHDKVWCG YRNKIYVVQFKAMKIESFDAHPRKESQVRQ LAWVGDGVWSIRLDSTLRLYHAHTYQHLQ							
HLSSETKAVLEEFGGFELELRGDVEMKGKG KVRTYWLLGERGSSTRG 804 2154 A 6585 2 3837 DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKPIEFEDSQEQEKK DLQTTVESLESQTRQLELLKAKNYADQISILEE RBAELKEYNALHQRHTEMIENYMPHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIPPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGSKATTPASTANSDVA TIPTDTPLKEENEGFVKYTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTFELL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALANIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEFELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKRSSIWQFFSILFSSSSNTTK KPEPPVILEYNAPTSHVTPSVKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVYLRPLDEKDTSMKLWCA VGVALSGGKTRDGSSVGASVFYKDVAGLD TEGSKQRSASQSSLDKLIQELKEQQKELKNQ EELSSLVWICTSTISATKVLIIDAVGPONILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSAETDSLLGGITVVGC SAEGVTGAATTSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGVYTPHVYTPIQLOQUPEDLSESG QVDKASLCGSMTSNSAETDSLLGGITVVGC SAEGVTGAATTSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGVYTPHVYTPIQLOQUPEDLSPYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRCCLHSIKLKD SLSIVIVKGTULVALADGTLAIFRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWSRLDSTLRLYHAHTYQHLQ			l				
KVRTYWLLGERGSSTRG		•					
MSERVSGLAGSTYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDOEHQVELELLRDIDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHIFEMIENYMEHLERT KLHQLSGSDQLESTAHSRIRKERFISLGIFPLP AGDGLLTPDAQKGGETFOSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPITDTPLKEEMEGYVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAHIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKFTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASREPAMOEKKSSIWQFFSLFSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCLASVFGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEENATEATEGNAGSAEDTVVJGS DAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHISIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ							
LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEKFIEFEDSÇEQEKK DLQTRYESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKKERISLGIPPL AGDGLLTPDAQKGGETIPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPPINSEISKHIEV QVAQETRNYSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKGFTRVE MARVLMEENYKYERIMELQEAVR WTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNIKYNAPTISHYTPSVKKRSSTLSQLP GDKSKAFPILSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPYVYYRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASYYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHASTKVLIIDAVQPGNILDS FTVCNSHYLLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPYMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDITVDIS QTGYYTEHVPTDPLGNUPIEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGITVLVALADGTLAIFHRGVDGQW DLSNYHLJDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQFKAMKIEKSFDAHPRKESQVRQ	804	2154	Α	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
QLTTQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKEEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSOPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNITESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAFAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSKIFSSSSNITTK KPEPPVNILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATTEATEGNAGSAEDTVDIS QTGYYTEHVPTDPLGIVQIPEDLSPVYQSSND SDAYKGDLSVLPNEGDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKLIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTIRLYHAHTTYQHLQ			ł				MSERVSGLAGSIYREFERLIVRYDEEVVKELIP
DLOTRVESLESOTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGL1TPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGSKATTTPASTANSDVA TIPPTDTPLKEENEGFVKVYTDAPNKSISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITTK KPEPPVNLK YNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCTASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPYMDKPPEMAEN SEVDENVPTABEAATEATGNAGSAEDTVOIDS QTGVYTEHVPTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRESQVRQ LAWVGDGVWYSIRLDSTLRLYHAHTTYQHLQ				ļ			LVVAVLENLDSVFAQDQEHQVELELLRDDNE
REAELKKEYNALHORHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIKRERISLGIPLP AGDGLLTPDAQKGGETPGSEQWKFQELSOPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPITDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPYYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVQPKAMKIEKSFDAHPRKESQVRQ			l				QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGSSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNYSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTTSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEVATEATEGNAGSAEDTVDIS QTGVYTEHVTIPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHIDKVWCG YRNKIYVVQPKAMKEKSFDAHPRKESQVCG YRNKIYVVQPKAMKEKSFDAHPRKESQVCG YRNKIYVVQPKAMKEKSFDAHPRKESQVCG							
AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDATNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWGFFSRLFSSSSNTIT KPEPPVNILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEAATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVQPKAMKEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							REAELKKEYNALHQRHTEMIHNYMEHLERT
SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPITDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRITRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSNITIK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTING QGENKMKNLPVPYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSSEG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNYSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELBAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCTASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QIGVYTEHVPTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGYDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKLEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRTTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVNDIS QTGVYTEHVPTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGYDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGIVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVYQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					}		
FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					İ		
LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSIPTAGRKFITVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCLASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEGNAGSAEDTVDIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					ĺ		
RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNITK KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNILK YNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNIKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVVDIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							MARIA AERMONKERI MELOPANDUTTAN
KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ]		
GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ						İ	
VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ	ļ						
VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVPTTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ						1	
TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTVDIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					\	J	
EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ						į	
SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					. }	ļ	
SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS QTGVYTEHVPTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ			•				•
QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					1	l	
SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					}		
MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					ļ		SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					l		MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					ļ		SILSIVHVKGIVLVALADGTLAIFHRGVDGQW
YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ							
LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ					!		
DUDIEDINION A OTOM CEOTION CONTRACTOR					i		
I JUVDIEFY VSKMLGIGKLGFSFVRITALMVSC					1		DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NRLWVGTGNGVIISIPLTETVILHQGRLLGLR ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
805	2155	A	6605	469	2602	FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ DSGLYACVIRNSTYCMKVSISLTVGENDTGL CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT REPEILWYKECRTKTWRPSIVFKRDTLLIREV REDDIGNYTCELKYGGFVVRRTTELTVTAPL TDKPPKLLYPMESKLTIQETQLGDSANLTCRA FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE SDIKILKEHLGEQEVSISLIVDSVEEGDLGNYS CYVENGNGRRHASVLLHKRELMYTVELAGG LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA EELDGDNKDYDAYLSYTKVDPDQWNQETGE EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF ELETRLRNMLVTGEIKVILIECSELRGIMNYQE VEALKHTIKLLTVIK WHGPKCNKLNSKFWKR LQYEMPFKRIEPITHEQALDVSEQGPFGELQT VSAISMAAATSTALATAHPDLRSTFHNTYHS QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA ILPLLPRETSISSVIW
806	2156	A	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL IQNKYFGDVDIPRAKVVRVCQALMDYKVFE AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD SQLGKENKLYSPARYADALFKSSDIRSASLED LWENLSLKPANSPHVNISTTLSPQVINEVWQE ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT LUHKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKK\LLQGFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRI.TRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMI.I.DLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\ AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LDLSSNPGLEVATGALGGLEASLEVLALQGN GLMVLQVDLPCFICLKRLNLAENRLSHLPAW TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR
						RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI NLIIILTFILVSAILLTTLAACCCVRRQKFNQQ YKA
	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG ARHRNNTEKKHPGGGESDASPEAGSGGGGV ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN AGSVGLALIVWIVTGFITVVGALCYAELGVNI PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLLTWVNCSSVRWATRVQDIFTAGKLL ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP DIGLVALAFLQGSFAYGGWNFLNYVTEELV DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD MYTLINYVGFINYLFYGVTVAGQIVLRWKKP DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI ELLTLVSQKMCVVVYPEVERGSGTEEANED MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV VLIDCVGVGLLIATLMWFISNKYLVKRQSRD YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL GYSVGLLFFSVALPFLKNTVILLYPFAPLILLYG LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811		A	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG AAASISQEPSLADIKARIQDSDEPDLSQNSITG EHSQLLDDGHKKARNAYLNISNYEEGDEYF DKNLALFEEEMDTRPKVSSLLNRMANYTNLT QGAKEHEEAENITEGKKKPTKTPQMGTFMG VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF AIVLICCCCTMLTAISMSAIATNGVVPAGGSY FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion FGHSKANGEPTWALLLTAAIAELGILIASLDL VAPILSMFFLMCYLFVNLACALQTLLRTPNW RPRFRYYHWALSFMGMSICLALMFISSWYYA IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ LVVAAKLREGISHLIQSCGLGGMKHNTVVM GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG GMLMLLPFLLKQHKVWRKCSIRFFTVAQLE DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
						DISAYTYERTLMMEQRSQMLRHMRLSKTER DREAQLVKDRNSMLRLISIGSDEDEE IETYQ EKVHMTWTKDKYMASRGQKAKSMEGFQDL LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL ERVLLVRGGGSEVITIYS
812	2162	Α	6628		640	AVCTMSEMAELSELYEESSDLQMDVMPGEG DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG CDEIIDRE
813	2163	Α	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC WQYRQLSALHRAPRPTRPDKARRLGYKAKQ GY/YYIYIGFVFAVIYRIRVRRGGRKRPVPKG ATYGKPVHHGVNQLKFARSLQSVAEERAGR HCGALRVLNSYWVGEDSTYKFFEVILIDFFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS RGLGKGHKFHHTIGGSRRAAWRRNTLQLH RYR
814	2164	A		201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN PLTKESIRQKEMESKRLRLLQEEDRRKKIARM GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF MHENLYFTNRKVNSVCWASLNHLDSHILLC LMGLAETPGCATLLPASLFVNSHPAGIDRPG\ MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR RVLLTNVVTGHRQSFGTNSDVLAQQFALMA PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

C020 ID	1 000 10	1 1 6 1	CDO	I Do to de la constante	I 8-11-1-1-1	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ĺ	ĺ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	!	ļ	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		ļ	ļ	sequence		nucleotide insertion
l	ĺ		1	i		RVFETLKDLKVLNLAYNKINKIADEAFYGLD
[NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
1		[ļ		QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
	ĺ	[Į		FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR
l	ļ		ļ	ļ	!	LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ
-	ļ	1			ł	TPSENPSLEQLFLGENMLQLAWETELCWDVF
	l		İ			EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR
	1	ł		ļ		GLSLNSNRLTVLSHNDLPANLEILDISRNQLL
1		ļ				APNPDVFVSLSVLDITHNKFICECELSTFINWL
	ĺ	ĺ	İ		ĺ	NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE GCDEEEVLKSLKFSLFIVCTVTLTLFLMTILTV
				ļ		TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY
ł			i			KYDAYLCFSSKDFTWVQNALLKHLDTQYSD
	ł					QNRFNLCFEERDFVPGENRP\ANIQDAIWNSR
1		İ				KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL
						NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ
l	ł	1	:			YLRWPEDLQDVGWFLHKLSQQILKKEKEKK
						KDNNIPLQTVATIS
816	2166	A	6646	1	3811	RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
010	2100	^	0010	•	3011	GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
						GGARLASLFGLDQAAAGHGNEFFQYTAPKOP
						KKGQGTAATGNQATPKTAPATMSTPTILVAT
		ŀ				AVHAYRYTNGQYVKQGKFGAAVLGNHTTR
						EYRILLYISQQQPVTVARIHVNFELMVRPNNY
						STFYDDQRQNWSIMFESEKAAVEFNKQVCIA
		1				KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE
						VAYTGWLFQNHVLGQVFDSTANKDKLLRLK
1					,	LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA
						CAVGSEGVIGWTQATDSILVFEVEVRRVKIA
	· ·					KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV
						VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD
] .			AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI
						EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ
1						MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA
						VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ
						PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
1						QHNTEIRMAVSKVADKMDHLMTKVEELQKH
						SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
						LKQEILEKSNRIEEQNDKISELIERNQRYVEQS
						NLMMEKRNNSLQTATENTQARVLHAEQEKA
		İ				KVTEELAAATAQVSHLQLKMTAHQKKETEL
						QMQLTESLKETDLLRGQLTKVQAKLSELQET
				\		SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL
						RVEKESLEKNLSERKKKSAQERSQAEEEIDEI
						RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS
				•		LVQAELQTQWEAKCEHLLASAKDEHLQQYQ
						EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA
				4		QITALTKQNEQHIKELEKNKSQMSGVEAAAS
J		l				DPSEKVKKIMNQVFQSLRREFELEESYNGRTI
						LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE
						EKAEERPRRPSQEQSASASSGQPQAPLNRERP
						ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR
						KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP
		İ				TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN
				,		PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
						ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE
						DELFKGATLKALRPKAQPEEEDEDEVSMKGR
						PPPTPLFGDDDDDDDDDDWLG
817	2167	Α	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG
317	210/	Λ	0049	03	10/3	TENSOSUNOSTINQ EDITOTRANQUEVINOLEU

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYDNLWDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPG\HLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLLTGLHACG\DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRPFGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG QALSVLETEDS
819		A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARNDRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\GRNEKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMTNLKDIGLYNLRNITRG\airieknadlcyl Stydwschldavsnyivgnkppkecgdlcp-GTMEEKPMCEKTTINNEYNYRCWTTNRCQK MCPSTCGKRACTENNECCHPECLGSCSAPDN DTACVACRHYYYAGVCVPACPPNTYRFEGW RCVDRDFCANILSAESSDSEGFVIHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFLK NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYKAVTLTM VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSQLIVKWNPPSLPNGNLSYYIVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPERKRRDVMQVA NTTMSSRSRNITAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMYYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKMGGKGLLPVRWMSPESL KDGVFTYYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS
821	2171	A	6691	106	825	ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC GRVLFRGCGVGHKGQVLMGTFILAQDWLSE
						SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK
822	2172	A .	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\IYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	Ą	6727	3	4063	PYLATLQLDSSLLIPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue ·	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	•	/=possible nucleotide deletion, \=possible
ļ		1		sequence		nucleotide insertion
						SSQPSQDGQESNVPSVGSLADPDYLNTPQMN
1		1			ł	TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP
						RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP
1 .				}		ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL
[1					SDSVMNIFKDRNFDSCCICACNMNIKGADVG
i i						LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL
į .						FLEDELDIFGKNSDIGQAAERRLMMCQSTFL
1						PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN
:			İ	ĺ		FLDYISSNNRQTLPCVSWSYDRVQADNNDY
ŧ						WTECFNALEQGRQYVDNPTGGKVDEALVRS
			'			ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP
1						FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH
j		ļ		ļ		KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT
						ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN
		Î				EALLEGAKTFFRDLSAVYEMCRLGQHKPICK
ł						VLRDGIMRVGKTVAQKLTDELVSEWFNQPW
						SGEENDNHSRLKLYAQVCRHHLAPYLATLOL
						DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN
						GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV
						PPVSSSASAPGISQISTTSSSGFSGSVGGONPST
	,					GGISADRTQGNIGCGGDTDPGQSSSQPSQDG
						QESVTERERIGIPTEPDSADSHAHPPAVVIYM
						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD
1						NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY
1						IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP
1						AASIEMTLKNPERPSPIQLYSPPFILAPIKDKOT
ļ i						ELGETFGEASQKYNVLFVGYCLSHDQRWLL
						ASCTDLHGELLETCVVNIALPNRSRRSKVSAR
						KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR
						LGHGELKDWSILLGECSLQTISKKLKDVCRM
						CGISAADSPSILSACLVAMEPQGSFVVMPDAV
						TMGSVFGRSTALNMQSSQLNTPQDASCTHIL
						VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL
í l						PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP
						SGIGVGSHFQHSRSQGERLLSREAPEELKQQP
						LALGYFVSTAKAENLPQWFWSSCPQAQNQC
						PLFLKASLHHHISVAQTDELLPARNSQRVPHP
						LDSKTTSDVLRFVLEOYNALSWLTCNPATOD
			0			RTSCLPVIIFVVLTQLYNAIMNIL
824	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
						RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG
[[GGGGGTIKRPGITGPTAATSPSGEPGNAASAP
						LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC
						ASLVFGRLQHRGGDRKRGLLGRSSGDAASD
						OPFRCRSGSTAGRLVKOMDFTEAYADTCSTV
1						GLAAREGNVKVLRKLLKKGRSVDVADNRG
						WMPIHEAAYHNSVECLQMLINADSSENYIKM
{		•				KTFEGFCALHLAASQGHWKIVQILLEAGADP
						NATTLEETTPLFLAVENGOIDVLRLLLOHGAN
						VNGSHSMCGWNSLHQASFQENAEIIKLLLRK
, 1						GANKECQDDFGITPLFVAAQYG\KLESL\SILIS
						SG\ANVNCQALDKATPLFIAAOEGHTKCVELL
					İ	LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI
						LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE
; l						DCLEILLRNGYSPDAQACLVFGFSSPVCMAFO
[[1	KDCEFFGIVNILLKYGAQINELHLAYCLKYEK
						FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA
i l			1			
i I						
						KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT
						LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G
825	2175	A	6735	277	1252	RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARAITDYLQ ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	2176		6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLYI ASDHTPLSFSVFERGFIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VGKGIMGVFTKPIGGAAELVSQTGYGILHGA GLSQLPKQRHQPSD\VHADQAPNSHVKYVW KMLQSLGRPEVHMALDVVLVRGSGQEHEGC LLLTSEVLFVVSVSEDTQQQAFPVTEIDCAQD SKQNNLLTVQLKQPRVACDVEVDGVRERLSE QQYNRLVDYITKTSCHLAPSCSSMQIPCPVVA AEPPPSTVKTYHYLVDPHFAQVFLSKFTMVK
827	2177	A	6748	2	1662	FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK ASGVSPTLWRPQAAATGLEMPSSGRALLDSP LDSGSLTSLDSSVFCSEGEGEPLALGDCFTVN VGGSRFVLSQQALSCFPHTRLGKLAVVVASY RRPGALAAVPSPLELCDDANPVDNEYFFDRS SQAFRYVLHYYRIGRLHVMEQLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DTEDQESQHESEQDFSQGPCPTVRQKLWNIL EKPGSSTAARIFGVISIIFVGVSIINMALMSAEL SWLDLQLLEILEYVCISWFTGEFVLRFLCVRD RCRFLRKVPNIIDLLAILPFYITLLVESLSG\SQT TQEL\ENVGAHCPGCLRLLRAL\RMLKAWGR HSTGLRSLGMTITQCYEEVGLLLFLSVGISIF STVEYFAEQSIPDTTFTSVPCAWWWATTSMT TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF
828	2178	,	6786	5672	1360	GTHPASSGPVPLPPAAVSAATREELGEPVFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVETIVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSPASSTPIPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNIKAPLTMASPAMLGNVESG GPPPTTASQPASVNIPGSSLPSSTPYTMPPEPTL SQNPLSIMMSRIMSKFAMPSISNPGYNHDAI

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KTVASSDDDSPPARSPNLPSMNNMPGMGINT QNPRISGPNPVVPMPILSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQVPFP HNGPSGQGSFPGGMGFPGEGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ YFPRGEVPGRKQPQGPGFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHIRMMSPAQST MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT HPGPVGSPGMMMSMQGMMGPNRTS
829	2179	А	6797	433	3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG GEREPRDRPSERPRLV
830	2180	A	6800	3	1911	LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC ARGDPASKSRSCGEVRQIYGAKGFSSSIDVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\ RELRLAT\RA\FVAAR\SFVQGLGVAS\IDVVR KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDTLTAKVIQGCGNPKVNPQGPGP EEKRRRGKLAPRERPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF ALLTLTVFLLLLADK VPETSLSVPIIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWQFVAMVVDRLFLW TFIIFTSVGTL\VIFLDATYHLPPPDFFP
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYWINPTLUS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV IMPCNWWVARMLGRV
833	2183	Α	6846	116	602	EAEGEQVCGAKCCGDAPHVENREEETARIGP GVMESKFERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFCYLMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVVV LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6855	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL
837	2187	A	6863	2	1615	PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEEFKVL LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA S\SVGKVVRRTQPVGTGPNL\YRKEYE\GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PNPRLEK,PVLLSEFPTK,IPRGELDMPVENGLR APVSAYQYALANGDVWKVHEVPDYSMAYG NPGVADATPPWSSYKEQSPQTLLELKRQRAA AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
838	2188	Á	6865	6291	739	EEMEEK VHGCCRIS AGPLEPR V QGAMAL QL WALTLLGLLGAGAS LRPRKLDFFRSEKELNHLAVDEASGVVYLGA VNALYQLDAKLQLEQQVATGPVLDNKKCTP PIEASQCHEAEMTDNVNQLLL VDPPRKRLVE CGQLLKGINCALRALSNISLRLFYEDGSGEKSF VASNDEGVATVGLVSSTGPGGDRVLFVGKG NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD KHPARNRTLLARMCREDPNYYSYLEMDLQC RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPSS IPVTPFQQDIIVAVTIQLLLRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQGPQAGGGTTLTITHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVNPOGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYYPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLIL.PLVIVPM VVVIAVSVYCYWRKSQQAERCTMKTLTET DLYCEPPEVQPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLIL.PLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPRRPVVEQALYQFSNLLNSKSFLINFIHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE RTKAITEIYLTRLLSVKGTLQQFVDNFFQSVL APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- ucnce	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, i=possible nucleotide deletion, \=possible nucleotide insertion HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL YAKEISTYKKMVEDYYKGIRQMVQVSDQDM NTHLAEISRAHTDSLNTLVALHQLYQYTQKY YDEIINALEEDPAAQKMQLAFRLQQIAAALE NKVTDL
839	2189	Α	6872		1485	RARRLALQCHVCVCALTPGEQSGRRLPGQT WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH EDQTDCSSLRDENNKENYPDAGALVEEHAPP SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP KSIFKAESGRSHGESQETEHVVSSQSECQVRA GTPAHESPQNNAFKCQETVRL\QPRIDQRTAT SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS TQSVLA\DGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRLLLYHITDGDNPTAVRHGCSL/F SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLLCTDV
840	2190	Α	.6873	2		FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN YLKQPVVKEKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA
						VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKRKEYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHE\APSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192	A	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSR\KTRHNST\DLPL
843	2193	A	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	Α	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195	Α	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
846	2106		6044	peptide sequence	2672	/=possible nucleotide deletion, \=possible nucleotide insertion
846	2196	Α	6944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP YLEVLETMDNKKGRRYEAVKWMVVFAIGV CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE AGSGITEGKCYLYARQVPGLVRLPTLLWKAL GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLLGATFNCLNKRLAKYRMRNVHP KPKLVRVLESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
847	2197	A	6951		1994	IVGIITRHNLTYEFLQARLRQHYQTI NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER LSSIEKIKQLREQVNDLFSRKFGEAIGVDFFVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEK WERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQVKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIK WLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2199	A	6999	963		LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLAVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL
850	2200	A	7001	1	1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG KVARRRVGATWLLHLAVADLLCCLSLPILAV PIARGGHWPYGAVGCRALPSIILLTMYASVLL LAALSADLCFLALGPAWCLRFS/GACGVQVA CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA SCHSALLCWAARRCRPLGTAIVVGFFVCWAP YHLLGLVLTVAAPNSALLARALRAEPLIVGL ALAHSCLNPMLFLYFGRAQLRRSLPAACHW ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	Λ	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI SRGVLVCDECCSVHRSLGRHISIVKHLRHSA WPPTLLQMVHTLASNGANSIWEHSLLDPAQV QSGPALKQTPKDKVVHPIKSEFIRAKYQMLAF VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG QTLQAELLVVYGADPGSPDVNGRTPIDYARQ AGHHELAERLVECQYELTDRLAFYLCGRKPD HKNGHYIIPQMADSLDLSELAKAAKKKLQAL SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD NLELSLRSQSDLDDQHDYDSVASDEDTDQEP LRSTGATRSNRARSMDSSDLSDGAVTLQEYL ELKKALATSEAKVQQLMKVNSSLSDELRRLQ REIHKLQAENLQLRQPPGPVPTPPLPSERAEH TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK LSRHGSGADSDYENTQSGDPLLGLEGKRFLE LGKEEDFHPELESLDGDLDPGLPSTEDVILKT EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA KAAKQLVTITTREKKQ
852	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL TVKGLLKPSFSPRNYKALSEVQGWKQRMAA KELARQNMDLGFKLLKKLAFYNPGRNIFLSP LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID QRLQPQRKFLEDAKNFYSAETILTNFQNLEM AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM DERGTEGAAGTGAQTLPMETPLVVKIDKPYL LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV ARHVAAGAGHENKHGGSRRFPAGVAPRRAM ANVSKKVSWSGRDRDDEEAAPLLRRTARPG GGTPLLNGAGPGAARQSPRSALFRVGHMSSV ELDDELLEPLDMDPPHPFPKEIPHNEKLLSLKY ESLDYDNSENQLFLEEERRINHTAFRTVEIKR WVICALIGILTGLVACFIDIVVENLAGLKYRVI KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH
						VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI AMGVVGGVLGAVFNALNYWLTMFRIRYIHR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGSM\$YPLQLFCADGEYNSMAAAFFNTPEK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRREKVGVIVDVLSDTASNHNGF PVVEHADDTQPARLQGLILRSQLIVLLKHKVF VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR VFKLFRALGLRHLVVVDNRNQVVGLVTRKD LARYRLGKRGLEELSLAQTGPKAQATAEGRV AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL RRQVQQEAERRGAEQEAGLRLAKLTDLLQ QEEQGREVACGALQKNQEDSSRRVDLEVAR
854		A	7037	139	2604	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDQDPDPQPRY TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFI DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	Α	7088	320	2417	LRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLG\LP WALIFFSFASGTFQLVVLYLFSITSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	Α	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

NO. of No. of No. of No. of No. of No. of No. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Sequence							
1914 ng to first amino acid residue of peptitide sequence 1914 ng to first amino acid residue of peptitide sequence 1915 ng to first amino acid residue of peptitide sequence 1915 ng to first acid sequence 191			}	1			
Peptide of peptide sequence		-	}	914		acid residue	Q=Glutamine, R=Arginine, S=Serine,
Peptide			1	ĺ	amino acid		T=Threonine, V=Valine, W=Tryptophan,
	1					sequence	
BPPEEDQGEERRILWAMPNHQYLLGFEEDQ DHIYHPY-GSSGHHLOFPYPRPHLGLGFSLP CPS							
B61 2211 A 7161 1220 1003 NYVCTIAFTEKMGFFLSLGLVLLFVLFLDG	<u> </u>				sequence		
S61 2211 A 7161 1220 1603 NYVCTIAF*PEKKMGF*LSLSCLVLLFVLFLDCL LTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ LTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ LXYYHITMGIYKTGKKVLL*KSSMSNRFSVF			İ	!			
S61 2211 A 7161 1220 1603 NYVCTIAF*EKKMGF*LSJ.SCL.VLLFVLE.DOC LTTTRIMFICTY_RASVCLSLNTLLSPNCL LTTTRIMFICTY_RASVCLSLNTLLSPNCL LTTTRIMFICTY_RASVCLSLNTLLSPNCL LTTTRIMFICTY_RASVCLSLNTLLSPNCL KSAMILQ	1		ł				,
	861	2211	A	7161	1220	1003	
\$62							
Second Second							KSAMILQ .
863 2213 A 7212 924 1273	862	2212	Α	7211	665	847	
L.VWDSPSCL.PATGFT*GLVL.VL.GGPDCT*WA RGQHEHKRMRP*SCRVTVALKKKKTDQ CIKPNYQSPPKECDYNILANSVA							
RGGHEHKRMRAPSCRVTVNLAKKKKKTDQ	863	2213	A	7212	924	1273	
CIKPNYQSPPKECDYNILANSVA			,				
364 2214 A 7214 845 1619 SDKGGKKADRKNHLRHAPPLLPHRYVERLH DPKYPVDADHVQGQDPGRAAHDIIIGDVTE KVSKDPLAPDEVQDTDEGIIDRHGHREVQGR HGHDQEEVAYEERACEGGKFATVEVTDKPV DEALREAMPKVAKYAGDFGGVAKEAD YDAQATRI.RALEGTATYRGDIYFCTGYDPP PAPSDKSVKIEREGGITVNGGIGMGMTV PISFAVFFNEDGSLQKKI.KVWFRIPNQFGSDP PAPSDKSVKIEREGGITVNGGIGMGTV PISFAVFFNEDGSLQKKI.KVWFRIPNQFGSDP PAPSDKSVKIEREGGITVNGGIGMGTV PISFAVFNEDGSLQKKI.KVWFRIPNQFGSDP PAPSDKSVKIEREGGITVNGGIGMGTV PISFAVFNEDGSLQKKI.KVWFRIPNQFGSDP PAPSDKSVKIEREGGITVNGGIGMGTV PISFAVFNGGTGT TYPKDP GROWN GREAT TYPKDP			ĺ .				
DPKYPVDADHVQGQPGRAAHDIIGEDVTE	864	2214	A	7214	845	1619	
KVSKDPLAPDEVGDTDEGHDRIGHREVGOR HOHODGEVAYERRACEGKRATVEVTDKPV					J.2		
DEALREAMPK VAKY AGGTNIDKGIGMGMTV	1						
PISFAYFPNEDGSLQKKLKVWFRIPNQFQSDP	}					J	
PAPSDKSVKIERERGITVYSMOFGGYAKEAD							
See						j	
MKPYGRRNEIWLLKT							
865 2215 A 7246 559 682							
LANMAKPILY	865	2215	A	7246	559	682	
DLKKSDFSTRWQKQRCPVVKSKCRENASPFF FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCOPENNIUCYKNNCCYGFPESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTINGSWQWEDGSILSPNLLT IIEMQKQDCALYASSFKGYIENCSTPNITYICM QRTV 867 2217 A 7288 151 396 SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI MPFFQTLWIMANRFCSIFTTINVANNCWW TPYHCWLSVVVCRCESHGI PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSMVW*VSESP*LGGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP 868 2218 A 7298 3 272 PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSMVESD*PEGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP 869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLHPINSVNSMAR FLTLCTWLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPESHLLA KRYGGFMKKVGGFMKKDAEEDDSLANSSDLL KELLFTGINRERSHHQDGSDNEEVSKRYGG FMRGLKRSPGFMKFVGGFMKRVG FMRGLKRSPGKRYGGFMKRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMKRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRFVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRFVG PGKSGRGCGGRAPRPEAMENGAVYSPT TEEDPGPARGRSGLAAYFFMGRIPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTFPYERVDTTKVKSSD FYITLGTGCYLLASIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA SCAMCSGLL*LLLPIWSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRGFRERAAPRR				,			
FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCGPCPKNWICYKNNCYGFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYYCM QRTV 867 2217	866	2216	Α	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL
QIPLTESYCGPCPKNWICYKINNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLKLV KSYHWMGLVHIPTINGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYICM QRTV					,		
WYESQASCMSQNASLLKYYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT	1						
RSYHWMGLVHIPTNGSWQWEDGSILSPNLLT	 					*	
BEMQKGDCALYASSFKGYIENCSTPNTYICM QRTV							
QRTV	1						
MPFFQTLWLMNANRFCSIFTTTNVANNCWW TPYHCWLSVVVCRCESHGI 868 2218 A 7298 3 272 PDTVIGGRGSGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP 869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKMDEL YPMEPEEA NGSEILAKRYGGFMKKMDEL YPMEPEEA NGSEILAKRYGGFMKKMDEL SLANSDOLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG PMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 I018 EIHQRLTERTQFLDESRKNPNS*QANILRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFTIMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR		•					
### TPYHCWLSVVVCRCESHGI ### TPYHCWLSVVVCRCESHGI ### TPYHCWLSVVVCRCESHGI ### TPYHCWLSVVVCRCESHGI ### TPYHCWLSVVVCRCESHGI ### PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP ### KGSCPAGGSRMVSESD*EGRGC*ASYPCAC*	867	2217	Α	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI
868 2218 A 7298 3 272 PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGGG*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP 869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGILHPNSVNSMAR FLITLCTWILLILGPGILATIVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKMDELSVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLIPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP 869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKDAEEDDSLANSSDLL KELLETGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
AGS*WR*GSRPAGRGTPPRSLSHARPP 869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWILLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHILLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLETGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGEGAESGGSRGEGFGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILILVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFLASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	868	2218	Α	7298	3	272	
869 2219 A 7332 1223 332 PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHILLA KRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDIL KELI-FTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGEGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	869	2219	A	7332	1223	332	
YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	"					·	
KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR		! !					YRLVRPADINFLACVMECEGKLPSLKIWETC
NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							KELLQLSKPELPQDGTSTLRENSKPEESHLLA
KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR				ks.			
FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF 870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
870 2220 A 7382 216 1018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
870 2220 A 7382 216 I018 EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGEGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPREAMENGAVYSPT TEEDPGPARGPRSGLAAVFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							•
AGQGRGREGAESGGSRGEGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTXVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	870	2220	Α	7382	216	1018	
TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							AGQGRGREGAESGGSRGEGPGSDGRLPATGD
KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	1						FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT
VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL
FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR	1 1		i				
GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
RAEALTEPLNA 871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR			-				
871 2221 A 7403 3 393 SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR							
DPGNMSFVKETVDKLLTGFRCFREREAAPRR	871	2221	A	7403	3	393	
ALRGAALPGESEAGDPESLRSSVNADWIQYS					ı		
							ALRGAALPGESEAGDPESLRSSVNADWIQYS

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
	<u> </u>	<u></u>		sequence		nucleotide insertion
						DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG PFIC
872	2222	Α	7413	1061	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
						CGLIKPCPANYFAFKICSGAANVVGPTMCFED RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ
						KAFDMYSGDVMHLVKFLKEIPGGALVLVAS YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
						SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE
873	2223	A	7429	2242	2394	GWPELLEMEGCMPPKPF ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ
						DHPGQHCETPSLLKIERKLF
874	2224	Α _	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISITE WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
						LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
	}	ł				AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
]	l .			MSSLNLDHWLKGAKREEWEPPPQSPALTHSP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
0.00	0005		7400		241	AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	Α	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG SERHEP*HGGVLFRLGPSAPPGKL
876	2226	Α	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	Α	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC
l		[ĺ			TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGREILELVQKGIPV
		Ì				TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
i		1				GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
						KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
						NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS
878	2228	A	7586	315	1232	ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
						RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
		1				QTFRGKENDTDLDLRYDTPEPYSEQDLWDW
						LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF
						RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
	6					AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
						YKLVQKVCPDYNYHSDTPYFPSG
879	2229	A	7605	479	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
880	2230	Α	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
						VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
						MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK
						CQSMGLLPTHASDWLAFIEPPERMEFSGGGL
881	2231	A	7615	291	1452	LL SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
						NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
						SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
					: 	CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI
	1					QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL
L	L	L		<u> </u>		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning .nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIIFTICFVPYHAFRFIYISS
882	2232	A	7617	67	379	QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST RQMALLKANKDLISAGLKEFSVLLNQQVFND
						PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	Α	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK AVATVGPISVAVGASHVFFQFYKKGKHLSS
884	2234	A	7638	2640	2861	APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP GLSCHTSHSG
885	2235	A	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	. ·	7702	242	1298	APSHRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGII.NPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707		2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFFLEQEADLIEAA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide		l	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
1	seq-	l			corresponding	
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i	Ì	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1			[amino acid	of-peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ľ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		Į		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			- !			EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
1					ļ	GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
		1			}	DESTKTKDQILTSRINAVERDLLEPSPADQLG
1 1					ļ	NGHRRTESEMSARIAKMSLSPSSPRHEDOLEV
						TREPARRLFLFGEEPSKLDQDVLAALECADV
1					!	DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
] .]		1			ļ	KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
		1				GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240		7711	360	269	
		A				RHMPVIPALWEAEVGGLLEPRSSRSAWATE
891	2241	Α	7721	61	1175	KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL
]					J	VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
}						HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN
						KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
					J	YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
1						AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
				'		WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
		•				KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR
					}	SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
						EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
l i						LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
{		}			ł	EAFKEWKEWIIKCLAVLQK
892	2242	Α	7723	2	1650	SAPTAPARPCRAERGSGGGMLALLAASVALA
""		``		_	.050	VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
!		ŀ				PAMPMQGGAQSPEEELRAAVLQLRETVVQQ
1 1						KETLASARAIRELTGKLARCEGLAGGKARGA
						GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
ļ						DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ
1		ĺ				LRETVVQQKETLASARAIRELTGKLARCEGL
])				AGGKARGAGATGKDTMGDLPRDPGHVVEQ
						LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
1						FREVLQQRLGELERQLLRKGAELEDEKSLLH
1						NETSAHRQKTESTLNALLQRVTELERGNSAF
1						KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
l 1						ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
						WGNNPIELLINDKVAQLPLFVSDGKWHHICV
1 1					•	TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
; l						KPGGVLILGQEQDTVGGRFDATQAFVGELSQ
1						FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
						NNVDVFGGASKWPVETCEERLLDL
893	2243	Α	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
1						DNYNDTSLVENHLCPATEGPLMASFKAVFVP
1			.			VAYSLIFLLGVIGNVLVLVILERHROTRSSTET
, [FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
1 1			1			LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
j 1			1			HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
	į					
	l l		1			LFAKVSQGHHNNSLPRCTFSQENQAETHAWF
i 1						TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
[]						QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
J J	J]		ļ	IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL
						GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
						CTGPASLCQLFPSWRRSSLSESENATSLTTF
894	2244	Α	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
			1	ĺ	Í	VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
				ļ		SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
				İ		VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
[1		D
895	2245	Α	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
						LWLSLFLHAGKEAPHCPRTRPL
896	2246	A	7754	1	372	SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL TKRGRQVCADPSEEWVQKYVSDLELSA
897	2247	Α	7761	1725		RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775 *	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	Α	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796		807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
903	2253	Α	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY RGDSPTDSQKDMIEIPLPPWQERTDESIETKR ARLLYESRKRGMLENCILLSLFAKEHLQHMT EKQLNLYDRLINEPSNDWDIYYWATEAKPAP EIFENEVMALLRDFAKNKNKEQRLRAPDLEY LFEKPR
904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG AGARLTGWTMNVFRILGDLSHLLAMILLLGK IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL WTFSIYLESVAILPQLFMISKTGEAETITTHYL FFLGLYRALYLANWIRRYQTENFYDQIAVVS GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL RSYSSI
905	2255	Α	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA QTEMVRTLERKLEAKMIKEESDYHDLESVVQ QVEQNLELMTKRAVKAENHVVKLKQEISLL QAQVSNFQRENEALRCGQGASLTVVKQNAD VALQNLRVVMNSAQASIEQLVSGAETLNLVA EILKSIDRISEVKDEEEDS
906	22.56	Ā	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG
907	2257	A	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG
908	2258	Α	7842	110		KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRHREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK
910	2260	A	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLELNKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNNTSLAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGFGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRELEEAPAPVLRPLSRPGSQECMDAAPG RLEFKDGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRTTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLWUPRYPVROWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT
911	2261	Ā	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ
912	2262	A	7891	1263	111	PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL
						EILVKEDRDSGVNFQPEDTCARLRCSLHASLL VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHK WIIAPKGFMANYCHGE CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD ECGCG
913	2263	A	7892	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN RLSNITRPFFSKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	Α	7914	3	967	TVAISTONANTOUCHART TVAISTONANTOUCHART VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK INCSWFIRANPGEIITISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCI.P ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	соптевропой	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ			peptide		/=possible nucleotide deletion, \=possible
_				sequence		nucleotide insertion .
						VILRETDEKLDGTGYGDYVKIYDGLEENPHK
		ļ		İ		LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV NAARGFNATYOVDGFCLPWEIPCGGNWGCY
		1				TEQORCDGYWHCPNGRDETNCTMCQKEEFP
		İ		İ		CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
	}		l	}		CQPGNFHCKNNRCVFESWVCDSQDDCGDGS
	•	ĺ	ļ			DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
			1			CTCKLYSLRMFERRSFETQLSRVEAELLRREA
		Ì				PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
	1					RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
		}				RSRHSGSLALVSADGDEVVPSQSTSREPERNH
						THRSLFSVESDDTDTENERRDMAGASGGVAA PLPQKVPPTTAVEATVGACASSSTQSTRGGH
		l				ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ
						GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG
	ĺ	ĺ	Ì			REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
			1			ASDQGQGLRQPYNATNPGVRPSNRDGPCERC
					•	GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
						GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQQTEDVLAA
1						GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
						VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
		l				RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS
1	-	i	i	Ì		FSSSWPSAQHLTPSVFNPW
921	2271	Α	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF
1		' '	133.	0.0	0.2	MMVYSFRALSFKESTWATFQHGGEATKSRSL
]	ļ			-02	SSTQ
922	2272	Α	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
	ļ	1				VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN
						WELVKI'N
923	2273	Α	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
			1	ļ		LLPLLLLPAGCRALEETLMDTKWVTSELAWT
Í			1			SHPESGWEEVSGYDEAMNPIRTYQVCNVRES SQNNWLRTGFIWRRDVQRVYVELKFTVRDC
						NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
			1		,	MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
])	}			FGPLSKAGFYLAFQDQGACMSLISVRAFYKK
		l				CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
1			}	•		AVEVSVPLKLYCNGDGEWMVPVGACTCATG
		ĺ				HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
}]	l	l			NSRTTSPAASICTCHNNFYRADSDSADSACTT
			ŀ			VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
					· ·	LLYNVICKKCHGAGGASACSRCDDNVEFVPR
		1				QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
1	•					GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
]		1	j		ļ ,	IASTVTSQMNSVQLDGLRPDARYVVQVRART
]		ĺ				VAGYGQYSRPAEFETTSERGSGAQQLQEQLP
						LIVGSATAGLVFVVAVVVIAIVCLRKORHGS
						DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA
			}			VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
			1			KQPGRREVFVAIKTLKVGYTERQRRDFLSEA
		l				SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
]			NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
]	1	Ī			KYLSEMNYVHRDLAARNILVNSNI.VCKVSDF
			[ļ		GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI
L	L		L	L		AYRKFTSASDVWSYGIVMWEVMSYGERPY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WDMSNQDVINAVEQDYRI.PPPMDCPTAI.HQ LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD WLDAIKMGRYKESFVSAGFASFDLVAQMTA EDLLRIGVTLAGHQKKILSSIQDMRLQMNQT LPVQV
924	2274	Α	7985	1	503	FRPRTKKATAMYLEHYLDSIENLPCELQRNF QLMRELDQRTEDKKAEIDILAAEYISTVKTLS PDQRVERLQKIQNAYSKCKEYSDDKVQLAM QTYEMVDKHIRRLDADLARFEADLKDKMEG SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS EEDTPKKKKHKGG
925	2275	Α	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR
926	2276	A	7996	925	582	SPIRCYCQHWPHCVHC GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS LNPKYSQIENFLSADMALKRKCLLSISDLDFW IWDAQPVGIMQTLQNLKKIPNPGCFWSQAFQI RDTQPILPLGGRYYITIRQ
927	2277	A	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV CLLLVTLALCCYQANAEFCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELMLPEKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279		8007	2	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVOPOGYSVVDFVRFECGEGEEAAETE
930	2280	Α	8008	3		NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP
931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	испсе		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	}	l		peptide	ſ	/=possible nucleotide deletion, \-possible
<u></u>				sequence		nucleotide insertion
						PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
1	!)	}		}	LIDYVRYMVENHGEDYKAMARDEKNYYQD
	ĺ		İ			TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
932	2282	A	0011	412	1	MEVE
932	2202	A	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
		1	ł			QGRGQIPIPCPWPPPPPPPPPPSPGPGCROFHQ
	i					SLEAKARHPASVREMRGKVKMRRALRRAPA
						STRASSROPNPK
933	2283	A	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
122		**	00.2			ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
						NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
			1			RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
	ļ	1				PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
						NASQLITQRAQVSLLIRRELTERAKDFSLILDD
	į	[VAITELSFSREYTAAVEAKQVAQQEAQRAQF
						LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
1		ł				SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
-004	0004		0000		000	DNLVLNLQDESFTRGSDSLIKGKK
934 ·	2284	Α	8023	255 .	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
						RLRKFRELHLMRNEARKLNHQEVVEEDKRL
						KLPANWEAKKARLEWELKEEEKKKECAARG EDYEKVKLLEISAEDAERWERKKKRKNPDLG
						FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
						KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
					,	KQIEKRDKYSRRRPYNDDADIDYINERNAKF
						NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
						QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
						SQHSSPAPMYSQTFHILVLG
936	2286	Α	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
1						FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
						IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
į			•			NEKSFONIRNWIRNIEEHASADVEKMILGNKC
1						DVNDKRQVSKERGEKLALDYGIKFMETSAK
						ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL
937	2287	Α	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	A	8052	675	.1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
1						PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
}						ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
ļ						LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
						GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
						VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	Α	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
						AEQLKWSAELARLGESIMDGKQGGMDGSKP
						AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
						IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
.						CLGGHLSCVKILLKHGAQVNGVTADWHTPL
1						FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
						PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTQAKN
						AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
						CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
						L
940	2290	Α	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
						ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

nucleotide sequence USSN 09495 uence unce colde sequence unce colde sequence unce e unce	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
1914 1915 1916	cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
Peptide		00,500	ļ				
Peptide		<u> </u>	1				
	l	İ	ł	}		sequence	
VVDTVMCPNMPNSVLLYTLSFYIFFVICMA ANSVVVWNIQAGATIGVPTHCVINIALADL WVVLTIPVWVSLVQHNQWPMGELTCKVTH LIPSINLPGSIPFLTCMSVDWTJSLYTTYTSS RKKMVRRVVCLVWLLAFCVSLPDTYYLKT VTSASNIETYCSS*PTSWBYSLISTYTTYTSS RKKMVRRVVCLVWLLAFCVSLPDTYYLKT VTSASNIETYCSS*PTSHSKEWLIGMELVSV VJGPAVPFSIJALWFYFLLARAISASSDOEKHISS RKIESYVVVTLVCWLPVHAVALLDIFSLIHTY PFTCRLEHALFTALHVTQCLSLVHCCVMPVL YSFRRNYRYTELMKAFIFKYSAKTGLTKLIDA SRVSETEYSALEQSTR 941 2291 A 8659 73 432 DDAGGLMTIVTSLLDLGVCAHHIIPTGSVVLPS PCCMFPVSRRPENRVVSYQLSSRSTCLKAGV IFTTKKGQQFCGDFKQEWQRYMKNLDAKQ KKASPRARAVAVKGPVGNTSHNOTTOLS RKASPRARAVAVKGPVGNTSHNOTTOLS RKASPRARAVAVKGPVGNTSHNOTTOLS VTSAVISQKSPRDICQGRISTLTQCVDSQVT MMFWYRQQFGGDFKQFSSTLMNSLGIGS VTSAVISQKSPRDICQGRISTLTQCVDSQVT MMFWYRQQFGSLTLIATANQGSEATYESGF VDKPTSSRNJLFSTLTVSMSPSEDSSYJLCSA GRQGTYEQYFGGTRLTVTEDLKNYFPPEVA VPEPSEAEISHTQKALTUCALGFYPPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS WWNNGKEVHSGVSTDFQPLKEQPALNDSRY VPEPSEAEISHTQKALTUCALGFYPHVELS VPEPSEAEISHTQKALTUCALGFYPHVELS VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEA VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCALGFT VPEPSEAEISHTQKALTUCAL							
ANSVVVWNIQAKTIGYDTHCYILINAIADI WVLITPVWVVSLVQHNQWMGELTCKVTH LIPSINLPGSIFFLTCMSVDRYLSITYFTNTPSS RKKMPRWVCLUVULAPCVSLPDTYVLKT VTSASNNETYCSFYPEHSIKEWLIGMELVSV VLGRAVPSILAVFYELARIASASOEKHSS RKIIPSYVVPLVCWLPYHVAVLDIPSILHYT FTCRLEHALFTALHYTQCLSLVHCCVNPVL SPRINNTRYELMKAFIFKYSAKTOLITKLIDA SRVSETEYSALEQSTK SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYSALESTS SRVSETEYS		 			sequence	<u> </u>	
			1	}			ANSVVVWVNIQAKTTGYDTHCYILNLALADL
RKKMVRRVVGILVMLLAFCVSLPDTYYLKU							
	İ					ł	
	ļ	1]			
		Ì	,				!
941 2291 A 8059 73 432 DMAGLMTIVTSILELGYTK 941 2291 A 8059 73 432 DMAGLMTIVTSILELGYCAHHIIPTGSVVLPS PCCMFFVSKRIPENRVVSYQLSSSTCLKAGV IPTTKKQQPCGDPKQBWQRYMKNLDAKQ IPTTKKQQPCGDPKQBWQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC 942 2292 A 8067 278 1262 GGIGEIKQRFSCLGRCLDPSLSVLMMISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQPGGSLTLLATANQSEATYESGF VIDKPPISRPNLTFSTLTVSIMSPEDSSIYLCSA GRQGTYEQYFGGFIRLLYTELKNPPPEVA VFEPSEARISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSCOSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPYTQVSAEAWGRADCFTS ESYQQVLSATTLYBULGKATLYAVLVSALV LMANVKRXDSRG RVEKOVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVGFYGLSE NDEWTQDRAKPYTQVSAEAWGRADCFTS ESYQQVLSATTLYBULGKATLYAVLVSALV LMANVKRXDSRG RVEKOVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVGFYGLSE NDEWTQDRAKPYTQUSAEAWGRADCFTS ESYQQCVLSATTLYBULGKATLYAVLVSALV LMANVKRXDSRG RVEKOVSTDPQPLKAPAPP PSDRKELLALPVACDPVISAWWQAASLI QGFGDKGDVFDEADESILAQREWGSMMQR RVEKOVSTDPGDLGGFNVGKGA EVILNYGRIRGTLSALLSWCHLHNNNSTLNK NNLLDAVQGCESVYLKHLAPPSPYLAVPAP PSDRKELLALPVACDPVISAWWQAASLI QGFGDKGDVFDEADESILAQREWGSMMQR RVEKOVSTDPGCESVYLKHLSTPPHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EPINKNCSKSHSGIDCSYVECCRTQEHAHSGK PKHMDFGTDSQF ESARWSRQLRRTILRISPFISCGRSHAFGGCK MAATSGTDPFVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKITKYDDQIVSFERNHETILARD DFE			ľ				RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI
SRYSETEYSALEQSTK		·	Į]		į	
941 2291 A 8059 73 432							
PCCMFFVSKRIPENRVSYQLSSRSTCLKAGV IFTTKKGQQFCQDPKQEWVQRYMKNLDAKQ IFTTKKGQQFCQDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC 942 2292 A 8067 278 1262 GGIGEIKQRPSCLGRCLDFSLSVLMYMISLGLGS VSAVISQKPSRDICQRGSLTLIACQVDSQVT MMFWYRQQFGQSLTLIATANQGSGATYESGF VIDKFPISPR-LTFSTLTYSMSPEDSSIYLCSA GRQGTYEQYFGGTRLTVTEDLKNVFPPEVA VFEPSEAEISHTYGKATLVCLATIGFYPDHVELS WWNNGKEVISGVSTDPQFLKEQFALMDSRY CLSSRLRVSATFWQNFRNHFRQVOYFYGLSE NDEWTQDRAKPYTQVSAEAWGRADCGFTS ESVQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAVFAP PSPRKELLALPVACDPVISASWSWQAASLI QGPGDKGDVFDEEADESILAQREWOSIMQR RVKEGYRDGIDGKAVTLQQFNQGVKKGA EVILNYGRLRGTLSALLSWCHLHANNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKDBAKDERLCENNA EFNKNCSKSHSGIDCSVVECCRTQEHAHSGK PKYHMDPGTISQF PSPRKELLALPVACDPVISHDAKASK PSPRKELLALPVACDPVISHDAKASK PSPRKELSKAPBGAKCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNETLRIDVLDPELK SESAKEKWRPCLKFRIGVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAEIARNRGGYNKA VYISVQDKEGKGKONGEKRADSGEENT KNGGGKGADSGEEKEGGINREDKTDKGGGK GKEADKRIFIKGSGKKAM VYISVQDKEGKKIFIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGKKHTIMKCSGTKKKEER ADILAYLKKATNE EGRGKFGGKLCRFLPYFINSNSAESRMDVLF VOHAPAPPLILIQGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSSBARDKULTK VOHAPAPPLILIQGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSSBARDRINLDK VAHAPAPPLILIGGSYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGEDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGDEYYQVVYY Y YTVTSYDDFSADTTIDSTSFSSDRINLDK VAHAPAPPLILIGGYEDEERIGGDEYY	941	2291	A	8059	73	432	
			l	0033	,,	132	
942 2292 A 8067 278 1262 GGIGEIKQRFSCLGRCLDFSLSVLMINISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLATANQGSEATYESGF VIDKFPISRPILTTSTLTVSNMSPEDSSTYLCSA GRQGTYEQYFGGFISHTVTEDLKNVPPPEVA VFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPYTQIVSAEA WGRADCGFTS ESYQQGVLSATILYEILIGKATLYALVSALV LMAMVKRKDSRG WKVVPATRONLPRSQLTGTHQHCQPREPKI TASSELRRRRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWQAASLI QQFQGKGVFDEEADSLAQREWQSMMQR QVKEGYRDGIDAGKAVTLQQGPQGVKKGA EVILNYGRLRGTLSALLSWCHLINNINSTLINK INNLIDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDBRLCENNA EFINKNCSKSHSGIDCSYVECCETQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 I 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYON DPDIEMAWAMRAMQHAEVYXLISSVDPQF KLTIKVDQJYSEPGRTERLBIVLDPEELK SESAKEK WRPFCLKFNGIVEDPNYGTLLRLD CSQGYTEENTIFAPRQFFAEIARNRGGEKAN VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEEGINRDKTDKGGEK GRADKENKSGEKAM VYISVQDKEGEKGVNIGGEKRADSGEENT KNGGEKGADSGEKEEGINRDKTDKGGEK GRATLLRSASSAARKAAEAEQVULPILHRVI SADRRVLGLREWGRPASERCSLCQRLKREL NMGDVEKGKKIFINKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSTAANKNKGIIW GEDTLMEVLEPPKRYUFGTKMIFVGIKKKEER ADIIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEBEERLGEDEYYQVVYY YTYTYFSVDDFSADFTIDVSIFSSEDRLNRLDK DIFFALETTISLETARADFFYTYKPYTTFPQ SPRSAAMFCPVLRSPPLPPVRVPLFRWGCISC KVGRRLLMTLWMGVWQEEIGR GGGSSPRELAGAAGLTVTSQAVAARQQPSF							IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ
943 2293 A 8070 1 879 MYKVPATRONLPRSQLATIFISTOGRESSITYLCSA GRQGTYEQYFGGTRLTVTEDLKNVFPPEVA VPESSEASISHTOKATLYCLATGFYPDHYVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVGFYGLSE NDEWTQDRAKPYTQIVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKKDSRG UMAVVPATRONLPRSQLTGTHQHCQPREPKI TASSRLRRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QCPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHINNISTLINK INNLLDAVGQCEEVVLKHLKSITPSHVVDLL DSIEDMDLCHVVPAEKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 1 797 ESARWSQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVHALLSLPAESYGN DPDIEMA WAMRAMQHEAYYYKLISSVDPOP LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEK WRPPCLKFNGIVEDRYGTLLRLD CSQGYTEENTIFAPRIOFFAIELARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKENKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRFSERCSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTOGAPGYSTAANKNKGIUW GEDTLMEVLLEPKRYVYDGTKMIFVGIKKKEER ADIJAYLKKATNE 946 2296 A 8081 42 590 EGRGKFGGKLCNFLFYFHSNSAESERMDVLF VAIFAVPLILGQEFSDEERLGEDESYYQVVYY YTVTPSYDDFSADFTIDVSIFESEBRLNRLDK DITEALETTISLETARADHFVPTVXPVTTEPQ SPRSEAMPCPVLRSPPLPPVRVPLFRWGCISC KKVGRRLLATLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTUTTSQAVAARRQQPSF							
MMFWYRQOFQSLTLIATANQGSEATYESGF VÜKPPISRPN.TFSTLTVSIMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VPEPSEAEISHTQKATLVCLATGFYPPHVELS WWVNOKEVHSGYDFQFLKEQPJALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPVTQVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG 943 2293 A 8070 1 879 MWYVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAYFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QCPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQGGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLIHNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF SEARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRFGLKFNGIVEDFNYGTLIRLD CSQGYTENTIFARPGTAEILARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGKGADSGEEKEGGINREDKTDKGGEK GKEADKEINKSGEKAM VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGKGADSGEEKEGGINFEDKTDKGGEK GKEADKEINKSGEKAM SOFTA SADRRYLGIREWGRASERSCSLOCRILKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GFNLHGLFGRKTQAPGYSYTAANKNKGIIW GEDTLMEYLENFKKYPGTKMIFVGIKKKEER ADJIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVELILGQEYEDEERLGEDEYYQVVYY YTVTTSYDDFSADFTIDYSIFFSEDRLINELDK DITFALIETTISLETTARADHPKPVTVKVPVTTEPQ SPRSEAMPCPVLRSPPLPPRVRPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSSPRELAGAAGLTVTSQAVAARRQQPSF	942	2292	A	8067	278	1262	
VIDKPPISRPNLTFSTLTVSIMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VPEPSEAEISHTQKATLVCLATGFYPDHVELS WWWNOKEVHSGYSTDPQPLKEQPANDBSY CLSSRLRVSATFWQNPRNHFRCQVFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS ESYQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRDSSG 943 2293 A 8070 1 879 MYKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACOPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKVTLQQGFTQGYKKGA EVILNYGRLRGTLSALLSWCHLHINNISTLINK NNLLDAVGQCEEYVLKHLKSITPPSHVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFINKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSGF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKADSGEEKEGGNREDKTDKGGEK GKEADKENKSGEKAM 945 2295 A 8074 2 505 GAATLLRSSSAARKAAEARQVWLHLHRYL SADRRVLGILRWGRASERCELCQRI KREL NMGDVEKGKIFIMKCSQCHTVEKGGKIKT GPNLHGLFGRKTQQAPGYSYTAANKNKGIIW GEDTLMEYLENKKYPGTKMIFVOIKKKEER ADIJAYLKKATNE 946 2296 A 8081 42 590 EGRGKFGGKLGVFFFFISNSAESRMDVLF VAIFAVPLILQQEYEDEERLGEDEYYQVVYY YTVTSYDDFSADFTIDYSIFFSEDRLINRLDK DITFAILETTISLETTARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPPLPPVR VPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAGGTVTSQAVAARRQQPSF							
GRGGTYEQYFGPGTRI.TVTEDLKNVPPPEVA			ļ				
WWYNGKEVHSÖVSTDPQPLKEQPALNDSRY							GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA
CLSSRLRVSATFWQNPRNIFFRCQVQFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS STYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRDSRG 943 2293 A 8070 1 879 MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNISTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGFUEDFNYGTLLRLD CSQGYTEENTIFAPRIGFARIEARNEGYNKA VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKERNKSGEKAM SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADIJAYLKATOL GFNLHGLFTIISLETTIISLETTIISLETTISLETAADHFKPYTVKPYTTEPQ SFRSEAMPCPVLRSPPILPYRVPLFRWGCISC KKVGRILLMTLWMGVWQEEGGR GGGSSFRELAGAGLTVTSQAVAARRQQPSF							,
NDEWTQDRAKPYTQIVSAEAWGRADGGFTS]		ļ				· · ·
SEYQQQ'LSATILYE'ILIGKATLYAVLVSALV							, , ,
943 2293 A 8070 1 879 MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRIGTISALLSWCHLINNNSTLINK INNLLDAVQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFFNKNCSSKHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTIKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEGINREDKTDKGGEK GKEADKEINKSGEKAM VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEGINREDKTDKGGEK GKEADKEINKSGEKAM SATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERESLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKEER ADLIAYLKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVELILGQEYEDEERYGEDEYYQVVYY YTVTFSYDFSADFTIDSYSFESEDRLINLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGWGEIGR GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							` `
TASERLRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACOPVIASVMSWVQAASLI QGPGDKGPDEBADESLAQREWQSMMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVQQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 I 797 ESARWSRQLRRTILIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKADSGEEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATILRSASSAARKAAEAEQVWLHILHRYL SADRRVLGIREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADILAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTFSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVR VPLFRWGCISC KKVGRRLLMTLWMGGWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 1 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENITFAPRIOFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEGINREDKTDKGGEK GKADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLTSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENFKKYIPGTKMIFVGIKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTTSYDDFSADFTIDYSIFESEDRLINRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSSPRELAGAAGLTVTSQAVAARRQQPSF	943	2293	Α	8070	1	879	, , ,
QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 I 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEFVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEK WRPFCLKFNGIVEDFNYGTLIRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKENKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPVRVPLFRWGCISC KKYGRRLLMTLWMGWVQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
RVKEGYRDGIDAGKAVTLQQĞFNQĞYKKĞA EVİLNYĞRLRĞTLSALLSWCHLHNNNSTLINK INNLDAVĞQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLÇENNA EFNKNCSKSHSĞIDCSYVECCRTQEHAHSĞK PKPHMDFĞTDSQF 944 2294 A 8073 I 797 ESARWSRQLRRTLIRLSFPISCĞRSHAFĞĞÇK MAATSĞTDEPVSĞELVSVAHALSLPAESYĞN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEK WRPFCLKFNĞIVEDFNYĞTLLRLD CSQĞYTEENTIFAPRIQFFAIEIARNREĞYNKA VYISVQDKEĞEKĞVNNĞĞEKRADSĞEENT KNĞĞEKĞADSĞEEKEĞİNREDKTDKĞĞEK ĞKEADKEINKSĞEKAM 945 2295 A 8074 2 505 ĞAATILRSASSAARKAAEAEQVWLHLHRYL SADRRVLĞI,REWĞRPASERECSLCQRLKREL NMĞDVEKĞKKİFIMKCSQÇHTVEKĞĞKHKT ĞPNLHĞLFÖRKTĞQAPĞYSYTAANKNKĞIIW ĞEDTLMEYLENPKKYIPĞTKMİFVĞİKKKEER ADI.IAYLKKATNE 946 2296 A 8081 42 590 EĞRÇKFĞĞKLCNFLFYFİSNSAESRMDVLF VAIFAVPLILĞQEYEDEERLĞĞDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DİTEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPÇPVLRSPIPLPPVRVPLFRWĞCISC KKVĞRRLLMTLWMĞVWQEEIĞR 947 2297 A 8084 322 549 GĞĞSSPRĒLAĞAAĞLTVTSQAVAARRQQPSF							
INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 1 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKAADSGEEENT KNGGEKGADSGEEKEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGIREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTQQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF 944 2294 A 8073 I 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPELK SESAKEK WRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENIIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATILRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
BEFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF							
944 2294 A 8073 I 797 ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF	[
MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDPNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
DPDIEMAWAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATILLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF	944	2294	Α	8073	1	797	
LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGIREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADI.IAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF	ì						
CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA
GKEADKEINKSGEKAM 945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTQQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
945 2295 A 8074 2 505 GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF	945	2295	A	8074	2	505	
NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADI.IAYLKKATNE 946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT
946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
946 2296 A 8081 42 590 EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF	946	2296	A	8081	42	590	
YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF					_		
SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR 947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
947 2297 A 8084 322 549 GGGSSPRELAGAAGLTVTSQAVAARRQQPSF							
	947	2297	A	8084	322	549	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1							SRARAPAHSLRAALSLASSARSWGAVSRDRG

				F 2	· · · · · · · · · · · · · · · · · · ·	
	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 1				peptide		/=possible nucleotide deletion, \=possible
LL				sequence		nucleotide insertion
						PCPPAIMYQSSNKC
948	2298	В	8093	3905	846	MEPGEVKDRILENISLSVKKLQSYFAACEDEI
1						PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG
1						YWVLVVHFTRREAIKQIEVLQHVATNLGRSR
1 1	-					AWLYLALNENSLESYLRLFOENLGLLHKYYV
1 1						KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA
1						PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG
1 1						SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD
1 1						FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST
1						
1 1	ļ					ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS
1						DTTPVHTTSQEKEEAQALDPPDACTELEVIRV
1						TKKKKIGKKKKSRSDEEASPLHPACSQKKCA
	·]					KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE
1						GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL
1						SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP
1	-					GDAPERPPLCDFSEGLSAPMDFYRFTVESPST
1						VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE
1						GGGGEGQTPRPLEDTTREAQELEAQLSLVRE
						GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS
1 1		i				GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH
1	i					VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL
1 1						TDCYVYLLRKGATEKPYLVEEAVSYNELDY
						VSVGLDQQTVKLVCTNRRKQFLLDTADVAL
1 1						AEFFLASLKSAMIKGCREPPYPSILTDATMEK
1						LALAKFVAQESKCEASAVTVRFYGLVHWED
1 1		ſ				PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT
i 1]					SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP
1		ļ				LLSVNMGGEQCGGCRRANTTDRPHAFQVILS
1 1	ĺ					DPPCLELSAESEAEMAEWMQHLCQAVSKGVI
1						POGVAPSPCIPCCLVLTDDRLFTCHEDCOTSF
1						FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS
î l	Ì	i				QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY
	1					QVDLPHTAIQEASNKKKFEDALSLIHSAWQR
						SDSLCRGRASRDPWC*
949	2299	$\overline{\mathbf{A}}$	8095	9	2374	
" '		^	0073	′	2014	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEOO
1		l				
1	l	j	-	l	(EQFETELKYKMTINGKIAVLYLKKNKNLLAP
		j	1	l	1	GYTETYYNSTGKEITTSPQIMDDCYYQGHILN
		İ	- 1			EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI
	i	i	i		ĺ	HRDGQEHALFKYNPDEKNYDSTCGMDGVL
		I	1		İ	WAHDLQQNIALPATKLVKLKDRKVQEHEKY
į l	1			l		IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN
	ľ	ľ	1		ļ	YVNMLYKKLNTHVALVGMEIWTDKDKIKIT
	l	l	į			PNASFTLENFSKWRGSVLSRRKRHDIAQLITA
	l	į	1		Ì	TELAGTTVGLAFMSTMCSPYSVGVVQDHSD
	ł	l	ł	ł	ł	NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS
	ł	l	1			TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL
	l	l]			SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC
	ł	ł	l	i	ł	GTSEECTNICCDAKTCKIKATFQCALGECCEK
		i	ľ	1		CQFKKAGMVCRPAKDECDLPEMCNGKSGNC
		l				PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ
1	Į	}				CTELWGPGTEVADKSCYNRNEGGSKYGYCR
j l	j	ŀ				RVDDTLIPCKANDTMCGKLFCQGGSDNLPW
j	j	l	1			KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG
1 1	ļ	l	ł			DNKVCINAECVDIEKAYKSTNCSSKCKGHAV
į l	1	l	ŀ			CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG
	t	I	l			VLFPMAVIFVVVAMVIRHOSSREKOKKDORP
[]						, ~, , , , , , , , , , , , , , , , , ,
		- I	· · · · · · · · · · · · · · · · · · ·	ł	l	
						LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRILLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	Α	8133		1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNILELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFTYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DRYTTSRRLDPIPQI.KCVGGTAGCDSYTPKVI QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	Α	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS QAGSLV
958	2308	Α	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP. LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEGFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960	2310	A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

	T 2 - 2 - 2 -		1		r	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	uana		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
l delice			711	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
ł		l		sequence	ł	nucleotide insertion
			 	· · · · · · · · · · · · · · · · · · ·		DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
				[LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL
ļ	}	ļ	l	}	ļ	ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN
1	l					AGANLQNYGETSPDAISTNSEGAQENHDDLM
ļ						SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP
]]	ļ		J	DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE
	1	1			İ	ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
		ľ	1			EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT
	ŀ					KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
						SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL
		ł	ł			TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN
			l			FPEFSFDFTREQLMEENESLKQELAKAKMAL
961	2311	A	8172	1442	682	AEAHLEKDALLHHIKKMTVE
901	2311] A	81/2	1442	082	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI
		ĺ	l			ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG
						EVQVSDKERHTQLEQMFRDIATIVADKCVNP
	ŀ		ĺ			ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA
		1				LEVIKOLKEKMKIERAHMRLRFILPVNEGKKL
ł	ł	ł	1		1	KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI
		l	ł			DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	Α	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS
						VLRRMQKKYWKTKQVFIKATGKKEDEHLVA
	1					SDAELDAKLEVFHSVQETCTELLKIIEKYQLR
			<u> </u>			LNGMKS
963	2313	A	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ
						GMDLVWSAWYGKCVKGKGSLPLSAHGIVV
			ĺ			AWLSRAEWDQVTVYLFCDDHKLQRYALNRI
						TVWRSRSGNELPLAVASTADLIRCKLLDVTG
			1			GLGTDELRLLYGMALVRFVNLISERKTKFAK
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI
						NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDOEEDKNIVVDDITEOKPE
	ŀ		1			
	[•
			}			PQDDGKSTESDVKADGDSKGSEEVDSHCKK
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL
-						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA
-						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE
•						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT
-						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDDDDEEEDRMEVGPFSTGQESPTA
•						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQOEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQOEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDERMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS
						PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGGESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGTTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTTDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
964	2314	Α	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS
964	2314	A	8184	6	1393	PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF EPRRNFRDDSTRPRTRGRTRGRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTTDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQQPKKK RSFSLSFSLLSPSEMMALGAAGATRVFVAMV AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY
						QPYPCAEDEECGTDEYCASPTRGGDAGVQIC LACRKRRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVIPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHILREDDIHIYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nuceticle deletion, \=possible
968	2318	A	8211	sequence 2	409	nucleotide insertion ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP YHQYIVEDWQEKYKSQILNLEESKATIHENIG AAGFKMSP
969	2319	A	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR VPGRLLLLLPALCCLPGAARAAAAAAGAGN RAAVAVAVARADEAEAPFAGQNWLKSYGY LLPYDSRASALHSAKALQSAVSTMQQFYGIP VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR NKRYALTGQKWRQKHITYSIHNYTPKVGELD TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT KPNICDGNFNTVALFRGEMFVFKDRWFWRL RNNRVQEGYPMQIEQFWKGLPARIDAAYER ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG ELGSCLPREGIDTALR WEPVGKTYFFKGERY WRYSEERRATDPGYPKPITVWKGIPQAPQGA FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY PRNILRDWMGCNQKEVERRKERRLPQDDVDI MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN YELLEKEVEKRKVALQEAKLKAKGLNPDGTP ALSTLGGFSPASKPSSPREVKAEEKSPISINVK TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR SASRSRSRTSRSRSHTPRRHYNNRRSRSGTY SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA KKHRHERGHHRDRRERSRSFERSHKSKHHGG SRSGHGRHRR
971	2321		8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN VEVEQESFFEGKNMALFEEEMDSNPMVSSLL NKLANYTNLSQGVVEHEEDEESRREAKAPR MGTFIGVYLPCLQNILGVILFRLTWIVGVAG VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA GAMYILGTIEIFLTYJSPGAAIFQAEAAGGEAA AMLHNMRVYGTCTLVLMALVVFVGVKYVN KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV CLLGNRTLSRRSFDACVKAYGIHNNSATSAL WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA ASGVFLENLWSTYAHAGAFVEKKGVPSVPV AESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVTVIGSFFSTCGAGLQTLTGAPRLL QAIARDGIVPFLQVFGHGKANGEPTWALLLT VLICETGILIASLDSVAPILSMFFLMCYLFVNL ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL CLALMFICSWYYALSAMLIAGCIYKYIEYRG AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

972 2322 A 8224 701 246 TSRRVTMKFNPFVISDRSKNRKRIHFNAPSHY RRKIMSSPLSKELRQKYNSMPTRADDEVQ VVRGHYKGQQIGKYVQVYRKKYVIYIERVQ REKANGTTYHVGIHPSKVVTITILKIDKDRKKI LERKAKSRQVGKEKGKYKEELIEEMQE 973 2323 A 8237 873 4610 GCPHAGGKGRVPTGGLTGGRTWSPSAPRSC PROPTPAGAMDKLIPPSMRKRLVSLFQQVG AKAWIMDEEDABEEGAGGRQDPSRSIRIR, PLPSPSPSAAAGGTESRSSA GAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGSGGTFP GLAAFERFGASAQPAASAPPPQQPPQFASAS CEQPSVDTAIKVEGGAAGDQLIPEAVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIHPYSDFFFYWDLTML LLMYGNLIIPWGTFFFABGAAGAAGDQLIPEAVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIHPYSDFFFYWDLTML LLMYGNLIIPWGTFFFABGAAGPAASDQLIPEAVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIHPYSDFFFYWDLTML LLMYGNLIIPWGTFFFABGYAATHWWSDF TFFLDLVLNFRTGIVEDNTEILDPQRIKMK YLKSWFMVDFISSIPVDYJFLIVETRIDSEVYK TARALRIVFFKILSLERLIRLSKLRIFYHHQWE EIFHMTYDLASAVVRIVNLIGMMLLCHWDG CLQFLVPMLQDFPDDCWSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMVGATCYAMPIGHATALIQSLDSS RQQCEKYKQVEQYMSPHKLPDTRQRIHD YYYEHRYQQKMFDEESILGELSEPLREEINFNC RKLVASMPLFANADPNFYTSMLTKLREVFQ PGGYIIREGTIGKKMYFIQHGVVSVLYKGNKE TKLADGSYFGEICLLTRGRTASVRADTYCR LYSLSVDNFNEVLEEYFMMRRAFETVALDRL DRIGKKNSLILHKVQHDLATHHPRLPAAIFR PPPGSGLGNLGAGGTPRHLKRLQSLIFSALGS ASPASSPSQVDTFSSSSFHIQQLAGFSAPAGLS PPLLPSSSSSPFIQQLAAAATTSVALATTHHPRLPAAIFR PPPGSGLGNLGAGGTPRHLKRLQSLIFSALGS ASPASSPSQVDTFSSSSFHIQQLAGFSAPAGLS PPLLPSSSSSPPPGAGGSFSSSFSPLOLAGFSPCHSFPGPEP PSLVAGASGASPVGTFPSSSSFSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
973 2323 A 8237 873 4610 GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PROPTPAPGAMDKLPPSMRKILYSLPQQVG AKAWIMDEEEDAEEGAGGRQPSRSISILR PLPSPSPSAAAGTESRSSALGAADSECPARG AGKSTINDCCRFREGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPASPPPQQPPPGASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRGFGAMLQPOVNKSISLMFGSQKA VEREQERVKSAGPWIHPYSDFRFVWDLIML LLMYGNLIIPVGITFFKDENTTPWIVNIVVSW TFFLIDLVLNRFTGIVSUPIEILDPGRIKMK YLKSWFMVDFISSIPVDYIPLIVETRIDSEVYK TARALRIVRFTKLISLLRLLRLSRLIKYINQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDW WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHR YQGKMFDEESILGELSEPLREEINFNC RKLVASMPLFANADPNFVTSMLTKLRREVPQ PGDVIRREGTIGKKMYFQHGVVSVLTKGNKE TKLADGSYFGBICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSHQOLAGFSAPAGLS PLLPSSSSSPPGACGSPSAPTFSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQFGARSPQ AAQPSPAPFGARGGLGJLEHFLPPPSSRSPSS SPGQLGQPPGELSLGLATOPLSTPETPPRQPEP PSLVAGASGGASPVETPGGLSPGFISSFOPP RTFPSAPPRASGSHGSLLLPPASSPPPQVPQR RGTFPLTFORLTQDLKLISASQFALPQGAGGT LRRASPFISSGESMAAPLEPFRAGGGSGGGSS GGLGPPGRYGAIPGGHVTLPRKTSSGSLPPP SLLFGARTSSGGPHLTAGPGREFGARPEPV SLLFSNL 974 2324 A 8247 279 468 EYKQWERFLSCQNTNLGLGYGKPKGGGLL LVPVKDASRICSLITYLLSSHWNNLVVRSPVL	972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI
LVPVKDASRICSLTYLLGSHWNNLVVRSPVL							GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEIINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGGTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSSPPP RTFSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
	974	2324	A	8247	279	468	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	Į	09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
ucite			1 314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1])	į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
[\		peptide		/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
975	2325	Α	8249	62	1571	I.VALKNWKPKGTNIPAPQSPVFGEAVSGVYM
1	i	[ĺ	[MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK
	}			ŀ		EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA
Į	ļ					LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT
]	ļ	1				ILPQELQAWVQEHCPESAEEAVTLLEDLEREL
1	i I		İ	Į.		DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ
ļ				i		DPRKVRDCRLSTQHEESADEQKGSEAEGLKG
						DIISVIIANKPEASLERQCVNLENEKGTKPPLQ
Í		[ĺ			EAGSKKGRESVPTKPTPGERRYICAECGKAFS
Ì	į				· ·	NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS
						NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK
}	ľ	l		1		HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR
						IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT
ł	İ					GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK
)	ļ	1		j	<u>}</u>	PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA
976	2326	A -	8257	298	7086	P GNMACWPOLRLLLWKNLTFRROTCOLLLE
310	2320] ^	8237	256	7000	VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM
						PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV
						VGNFNKSIVARLFSDARRLLLYSOKDTSMKD
	ļ					MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS
	İ	l	i			GFLYHNLSLPKSTVDKMLRADVILHKVFLQG
		}	<u> </u>			YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP
	ļ]				REKLAAAERVLRSNMDILKPILRTLNSTSPFPS
	į	ł				KELAEATKTLLHSLGTLAQELFSMRSWSDMR
				:		QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE
						TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK
]			PLLVGKILYTPDTPATRQVMAEVNKTFQELA
						VFHDLEGMWEELSPKIWTFMENSQEMDLVR
		Ī				MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL
				•		AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS
						RFMECVNLNKLEPIATEVWLINKSMELLDER
			· .			KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN
	į					VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP
	}		}			YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV
						IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI
		l				SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV
j		j				FLSVFAVVTILQCFLISTLFSRANLAAACGGII
						YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP
		1				VAFGFGCEYFALFEEQGIGVQWDNLFESPVE
1		I				PGOYGIPRPWYFPCTKSYWFGEESDEKSHPGS
[NOKRISEICMEEEPTHLKLGVSIONLVKVYRD
					İ	GMKVAVDGLALNFYEGQITSFLGHNGAGKT
						TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ
		1				NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS
]						EKHVKAEMEQMALDVGLPSSKLKSKTSQLS
<u> </u>		l i				GGMQRKLSVALAFVGGSKVVILDEPTAGVDP
]						YSRRGIWELLLKYRQGRTIILSTHHMDEADVL
						GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT
						LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS
						SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID
1						DRLSDLGISSYGISETTLEEIFLKVAEESGVDA
						ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA
		t l				ADPNDSDIDPESRETDLLSGMDGKGSYQVKG
]]						WKLTQQQFVALLWKRLLIARRSRKGFFAQIV
				·		

CEO ID	SEQ ID	Mot	I GEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	NO: of	Met hod	SEQ ID NO:		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	peptide	, 1100	in NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1	1	ļ	USSN	location	corresponding	l=Isoleucine, K=Lvsine, L=Leucine,
eotide	seq- uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	dence	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	j	į	314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]		residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	ļ	i		sequence	
ŀ		Ì		peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
i	i	l	Ì	i		LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
		ŀ				NEQYTFVSNDAPEDTGTLELLNALTKDPGFG
						TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
	I	1				DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
ļ]		CPPGAGGLPPPQRKQNTADILQDLTGRNISDY
	ļ	[1			LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS
		l	1	ì		NTQALPPSQEVNDATKQMKKHLKLAKDSSA
Ì	1	1	}			DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
1	1	1				HAISSFLNVINNAILRANLQKGENPSHYGITAF
1	1	l		ł		NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
	i		1			MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
1	i		i			YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY
		}		1		VSSTNLPVLALLLLLYGWSITPLMYPASFVFK
i	l	İ				IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN
ĺ	1	ĺ	[ĺ.		KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ
						AMADALERFGENRFVSPLSWDLVGRNLFAM
ľ	1	1	l			AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
}		ŀ				DEDEDVRRERQRILDGGGQNDILEIKELTKIY
1		!	l	Ī		RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
	1					SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
			ł			
		i	 	ļ		HONMGYCPOFDAITELLTGREHVEFFALLRG
1		l				VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY
ł	l	1	l	ľ		SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
	1	İ				PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
		ĺ				EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
		Ì				GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK
Ì	·		l			EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH
		ŀ	•			IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
	L					SLHKNQTVVDVAVLTSFLQDEKVKESYV
977	2327	Α	8260	3	1567	IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG
						YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
1	ı	ł		1		TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
						SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
	١.	İ	ŀ			YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS
		ŀ				LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM
	1					VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
ł	1	i :				MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
	1]		MYAYAGWFYLNFVTEEVENPEKTIPLAICISM
1	1		l	1		AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
			<u> </u>			FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV
1	1		1	'	l	SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
1	İ					LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA
	l		ł	i		VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC
l	1		l			LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
		1		1		WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
070	2220		8261		21/5	
978	2328	Α	8261	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL
						RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP
	1	l '	l			LADAASMSGVRAVRISIESACEKQVHEVGLD
	J		ļ			GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE
l	[EEAAGTEGDAQEWPGAGSSADQDDEEGVVK
						FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI
1			·			VRDKKFMTLDPVSQDALPPKQNPQTLQLISK
			[KKSLAGAAQILLKGAERLTKSVTENQENKLQ
1	Į į]	RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
						AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
						DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
1	i					LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
1						FAOLSREAVOIKSQVPHIVVKNOIISQPFPSLQ
1					i	LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
						HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- ucnce	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Mcthionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	Α	8289	2	1053	FVWNPRGGRKRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLFVLLMTLFFSGLILALLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	A	8315	1		GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFIYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF
984	2334	ı	8321		1243	DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY KAYLQDLVEGMDFQGPGES
		A		1		ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	Α	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RIRKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	Α	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIKKYDQMAIFHCLFWPSLTLI.GGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH QFKLCIMRRSKGRAEKS SSVVEFSALSVSMACLSPSOLOKFOODGFLVL
						EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS APGTSFLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906		MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRRWLGDPEHL
993	2343	A	8379	1		MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE DMVTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVNNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQITDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVILFIPQANPDDIDPT PTPTTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
995	2345	Α	8390	194	3421	PSSWDYRACLS AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		Ì		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	•		ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}]	1		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						DFLSMKQSPALAPEERCRAGSPKPVLRADD
i						NNMGNGCSQKLATANLLRFLLLVLIPCICALV
	1				ĺ	LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI
i		ļ				QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT
1]			,	TDASLPGDQSHRNTSACMNITHSQCQMLPYH
	ļ	!	ļ		!	ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY
ŀ					! !	QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE
1		1				AAKEGCESVLGMVNYSWPDFLRCSQFRNQT
		l				ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENL
1	İ	İ		!		FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC
						DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD
1		1				KSDEVNCSCHSOGLVECRNGQCIPSTFQCDG
1	!	1				DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP
		I				CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM
	Ì	l		!		NLPYNSTSYPNYFGHRTOKEASISWESSLFPA
1	[į				LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP
1		1				CRALCEHSKERCESVLGIVGLQWPEDTDCSQ.
1						FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ
		l				CVLASRRCDGQADCDDDSDEENCGCKERDL
		ŀ				WECPSNKQCLKHTVICDGFPDCPDYMDEKN
j						CSFCQDDELECANHACVSRDLWCDGEADCS
		1			İ	DSSDEWDCVTLSINVNSSSFLMVHRAATEHH
		-	1	ı		VCADGWQEILSQLACKQMGLGEPSVTKLIQE
		İ				QEKEPRWLTLHSNWESLNGTTLHELLVNGQS
]			CESRSKISLLCTKQDCGRRPAARMNKRILGGR
		ł				TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW
		ļ				VLTVAHCFEGRENAAVWKVVLGINNLDHPS
		ļ				VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
		ļ ·				DISETGYVRPVCLPNPEQWLEPDTYCYITGW
1						GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK TITTRMICAGYESGTVDSCMGDSGGPLVCEK
1						PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS
i						YFVEWIKROIYIOTFLLN
996	2346	A	8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSRS
""	2540	^	0372	199	3083	FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAHN
1		ŀ				RERNHPRVYONRDFRGHNRGYRRPYYFRGR
		İ				NRGFYPWGQYNRGGYGNYRSNWQNYRQAY
		l				SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS
						DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS
						KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK
		l '				ASESSKPWPDATYGTGSASRASAVSELSPRER
1		•			l i	SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST
		l				SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
1]	SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
1		•				YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE
1						KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
1				;		SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
						DKAKGRKESEFDDEPKFMSKVIGANKNQEEE
						KSGKWEGLVYAPPGKEKQRKTEELEEESFPE
1						RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
						AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS
						SFSITREAQVNVRMDSFDEDLARPSGLLAQER
						KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
						SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
						KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
1						HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
			•		İ	LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
						ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS
L		L	L	L		SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFOFRARGRGWG
				_		RGNYSGNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE DDESGTENREEKDNIQPTTE
997	2347	A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS KQENLESLEATGREEEGGQGPPVTTKGVLLA LLMAGLALQPGTALLCYSCKAQVSNEDCLQ VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL VSKKANQQLNFTEAKEACRLLGLSLAGKDQ VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE VFMETSTMSTETEPFVENKAAFKNEAAGFGG VPTALLVLALLFFGAAAGLGFCYVKRYVKAF PFTNKNQQKEMIETKVVKEEKANDSNPNEES KKTDKNPEESKSPSKTTMRCLEAEV
	2350	A	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT AAIFTADGNMISASTLMDILLMNDFKLVINKI AYDVQCPKREKPSNEHTAEMEHMKSLVHRL FTILHLEESQKKREHHLLEKIDHLKEQLQPLE QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWVYSWDIMEPVTYFITFANSMVFF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ HFDVQQYNKLKEDLAKAKESLKQARHSLCL QMQVEELNEKN
1001	2351	A	8410	1400		VGFWERPLRSSRWFRRSLRRWEMLARAARG TGALLLRGSLLASGRAPRRASSGLPRNTVVLF VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV LYLRIMDPYKASYGVEDPEYAVTQLAQTTM RSELGKLSLDKVFRERESLNASIVDAINQAAD CWGIRCLRYEIKDHVPPRVKESMQMQVEAE RKKRATVLESEGTRESAINVAEGKKQAQILAS EAEKAEQINQAAGEASAVLAKAKAKAEAIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA KDSNTILLPSNPGDVTSMVAQAMGVYGALT KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIIHNKHAG
1003	2353	Α	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLDTENIDEILNNADVALVNFYADWCRFSQM LHPIFEEASDVIKEEFPNENQVVFARVDCDQH SDIAQRYRISKYPTLKLFRNGMMKREYRGQ RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS KRNIIGYFEQKDSDNYRVFERVANILHDDCAF LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY LGAMTNFDVTYNWIODKCVPLVREITFENGE
						ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL ISEKGTINFLHADCDKFRHPLLHIQKTPADCP VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL HSGKLHREFHHGPDPTDTAPGEQAQDVASSP PESSFQKLAPSEYRYTLLRDRDEL
1004	2354	Α	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM ACAAARSPADQDRFICIYPAYLNNKKTIAEGR RIPISKAVENPTATEIQDVCSAVGLNVFLEKN KMYSREWNRDVQYRGRVRVQLKQEDGSLC LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA DQSLQQGEGSKKGKGKKKK
1005	2355	A	8453	90	530	QSHETKMQSGTHWRVLGLCLLSVGVWGQD GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF SELEQSGYYVCYPRGSKPEDANFYLYLRARG NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC QLCIFN
1007	2357	Α	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL . SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA FWWHNKGLALIFCILQSLALTWYSLSFIPFAR DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS DPRWGCVGPSMPTSTCLPGAVEASTTKASLP KCPVDSSLPTPEACFL
1009	2359	Α	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL TKLLVHSSLVGSILSALSALVGFIILSVKQATL NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT HDCGYEELLTS
1010	2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA HRVALCHLAGCQEQAAWYHTLQILFFLVSAY FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT GTLETQFTCPFCNHEKSCDVKMDRARNTGVI SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC RAGFQCQ
1012	2362	Α	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM RMKYGGQEFWADLNAMNVYETTEFDQLRR LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SVIRI.IEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKRKKEYMNRKMFGR DRIINERHLFHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI QYEEVSNTVSI
1013	2363	A	8488	2	\$17	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	A .	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2366	Α	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN
1017	2367	A	8513	54	1196	LERTPASADMAWTK YQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLI LVPMYYIPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEÄAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLRKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ
1020	2370	A	8530	2	1200	RREKGAP PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	ì	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	À	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAJVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	ng to first amino acid residue of peptide	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT
						WIVEFFANWSNDCQSFAPIYADLSLKYNCTG LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS
1025	2375	A	8546	2194	1707	TPTTVSDGENKKDK TVSFHKTMASLKCSTVVCVICLEKPKYRCPA
						CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS ALPTKTVKPVENKDDDDSIADFLNSDEEEDR VSLQNLKNLGESATLRSLLLNPHLRQLMVNL
ļ						DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES
1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL
						SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF
1028	2378	A	8569	20	963	VAAITSACHKYFEKAGLK KMAATLGPLGSWQQWRRCLSARDGSRRLLL
						LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY
					:	VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK
1029	2379	A	8572	1	578	RFY AAAASHRSRARSRPRRVSSGPAPRRAQSSAG
						RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM
	ļ					EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR
						AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWOSIKG
						THLTTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP
						SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL
					:	AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF
	 					CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG
			0.000			SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL
						FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
						VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW
1032	2382	A	8593	2558	961	HLDEVFLELKDGQQIPVFKLSGENGDEVKKE RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR
L			L			WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	derice	ŀ				Q=Glutamine, R=Arginine, S=Serine,
uence	ł		914	ng to first	acid residue	
1			!	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ	J	peptide		/=possible nucleotide deletion, \-possible
1	ì		ļ	sequence		nucleotide insertion
	l		i	· ·		QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
i						VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
[TFNNNDTVSFLEYRTFOFOPSKSHGSESDYIV
				ŀ		
	ŀ		İ	ŀ		MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
ł						GERAFMNRTVGEIMWGYKDPLVNLINKYFP
Į.	ŀ		ļ	ļ		GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI
İ	1		İ	ŀ		SRIHLVDKWNGLSKVDFWHSDQCNMINGTS
	İ					GQMWPPFMTPESSLEFYSPEACRSMKLMYKE
1			ļ	ĺ		SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
1			1			CLESGIONVSTCRFSAPLFLSHPHFLNADPVL
Ì			1	[AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
ļ			J			KLOLSLYMKSVAGIGOTGKIEPVVLPLLWFA
[1		ļ	ĺ		
l	,					ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
ŀ	İ			ĺ		LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
					,	GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	A	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
	1		1			FCLLLSLVSSSLVSLSLCPPLTOA
1034	2384	A	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYO
1034	2304	Λ.	6557	040	104	PMMOTIGOKYCMDPAVIAGVLSRKSPGDKIL
l	i					VNMGDRTSMVQDPGSQAPTSWISESQVFQTT
1						EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
1		İ				GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
ſ	[Í			F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
[SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
[i			SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI
						HVYKKNGVGKVGDQILLAIKGQKKKALIVG
<u> </u>	[HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
	}		!			1
	L					KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV
1						TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
ł	}		ļ			VAYRPISASVLSRPEASRTGEGSTVFNGAQNG
	1		1			VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
	ì		i			VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
1	j] .			AILGFALSEAMGLFCLMVAFLILFAM
1037	2387		8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
ן נטו	2301	Α	0013	_	4304	
1					}	GMVAHINNSRLKAKGVGQHDNAQNFGNQSF
1			[EELRAACLRKGELFEDPLFPAEPSSLGFKDLG
	1			•	,	PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
	[!			CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
						QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
ł	1		1			PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
				-		GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
	1		!	l		PONLLRLLRKAVERSSLMGCSIEVTSDSELES
ł			i l	!		
					ļ	MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
İ						RVRNPWGRIEWNGAWSDSAREWEEVASDIQ
1						MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL
						TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
1						RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
1						VVCTCLVALMQKNWRHARQQGAQLQTIGFV
	1		Į į			LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI
i				[İ	
						FTNSREVSSQLRLPPGEYIIPSTFEPHRDADFL
1						LRVFTEKHSESWELDEVNYAEQLQEEKVSED
1					İ	DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
1					1	MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
1						KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
1						LNSYEMRLVIEKAGIKLNNKVMQVLVARYA
						DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
1	ł					GHICLSLEOVLGEGWEGICRIAPACPSTPPPPS
		1	1	i		OTHERSTEW A FOR A ROTERIAL WEST LABOR.

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion
1038	2388	A	8621	3	1494	RSRMARAPLGVLLLGLLGRGVGKNEELRLY HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF GGIETLRVPSELVWLPEIVLENNIDGQFGVAY DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636		900	PGRERPGGGGARRRPQHLPALLPSERPDCATL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98		ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP GIKARITQRALDYGVQAGMKMIEQMIKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP
1042	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I≈Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l conce		}	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ		peptide	Sequence	/=possible nucleotide deletion, \=possible
		į.	1	sequence		nucleotide insertion
		 	 	Soquerree	·	TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT
1		1	l	i		LDMITSTDHVLEODFWICFTFYSVKEROI
1043	2393	A	8688	359	17	GLKTRAPATPTFQREVLGPAKQDMQRRCPRI
1043	2393	A .	0000	339	, ''	GLMTSLLKPIKRRWRDYKRWKSGGFTGESC
Į					1	
ł				1	1	HHADTLGDRGGLQGDHSELLQWQKRILRTE
		 			1100	GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS
[ĺ	([GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS
	1		1			YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL
						VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL
1					ł i	NLALADLLFALTLPIWAASKVNGWIFGTFLC
ŀ	Į.		1			KVVSLLKEVNFYSGILLLACISVDRYLAIVHA
				ĺ	,	TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR
Į.	ļ	İ			!	RTVYSSNVSPACYEDMGNNTANWRMLLRIL
1	l]			PQSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK
ŀ						HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM
ļ			ļ		i .	RTQVIQETCERRNHIDRALDATEILGILHSCLN
į.	1			1		PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS
ł	1	}		_		RPSFVGSSSGHTSTTL
1045	2395	Λ	8724	254	3184	FRANLAITVANRRGAQGGKMHTCCPPVTLEQ
ţ.	1					DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY
ł	ļ	l	l			YGEICDNACPCEEKDGILTVSCENRGIISLSEIS
1	ŀ					PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL
1	ł	1				HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL
j	,	}	;		j	ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF
						GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL
1						DLRGNRLKLLPYVGLLQHMDKVVELQLEEN
İ	l	İ		ļ		PWNCSCELISLKDWLDSISYSALVGDVVCETP
ſ	ĺ	ĺ				FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP
	J	j	ŀ			LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP
		ł				KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA
	ĺ	٠.				YOTKSPVPLECPTACSCNLQISDLGLNVNCQE
		İ	[RKIESIAELQPKPYNPKKMYLTENYIAVVRRT
1	[· '	1	[DLLEATGLDLLHLGNNRISMIQDRAFGDLTN
		l				LRRLYLNGNRIERLSPELFYGLOSLQYLFLQY
	İ					NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS
ì	!	[GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS
1	!					LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG
	ļ.	l				VLVDEVICKAPKKFAETDMRSIKSELLCPDYS
1		ł				DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA
		1				PASLGAGGASSVPLSVLILSLLLVFIMSVFVA
		Ì			i	AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN
						MOYSVYGGGGGTGGHPHAHVHHRGPALPK
1					1	VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN
		}				
	1	l				SVEDYKOLHELKVTYSSNHHLQQQQQPPPPP
1						QQPQQQPPPQLQLQPGEERRESHHLRSPAYS
]	VSTIEPREDLLSPVQDADRFYRGILEPDKHCST
						TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH
	1	1		'	Ì	QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE
		<u> </u>				YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	Α	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT
1						AMAGALVRKAADYVRSKDFRDYLMSTHFW
1						GPVANWGLPIAAINDMKKSPEIISGRMTFALC
		ł				CYSLTFMRFAYKVQPRNWLLFACHATNEVA
1						QLIQGGRLIKHEMTKTASA
1047	2397	Α	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP
ļ				1		AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK
						PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR
			لسننسا	نـــــــــــــــــــــــــــــــــــــ		

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VAVPNGQPPSAARYMPREVPPRFRCQQDHK VLLKRGQPPPSCMLLGGGAGPPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME NAGVNFVVSGREQAQIHNTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSQQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNK WGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNGMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPNSWNKQHQQQPPQQPPPPQ
						PEASGSWGGPPPPPGGNVRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEFQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF
1049	2399	A	8748	200	1387	ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA PFALSALLYGANNNLVIYLQRYMDPSTYQVL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL MAAGACYAAGGLQVPGNTLPSPPPAAAASP MPLHITPLGLLLILYCLISGLSSVYTELLMKR QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP GLLEGFSGWAALVVLSQALNGLLMSAVMKH
1050	2400	A	8758	3	1660	GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA FFLATLLIGLAMRLYYGSR WVSSMGFEELLEQVGGFGPFQLRNVALLALP RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS HQDVWLEAHLPREPDGTLSSCLRFAYPQALP NTTLGEERQSRGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFFA GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS
1051	2401	Α	8759	515	1625	TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE
1052	2402	Α	8763	1106	70	VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL
1053	2403	Α	8768	2	712	ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNVNVVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY
1054	2404	Α	8769	344	527	SSLMKVENMSSNQDGNDSDEFM REATTLACRNSCWVFSRCSLGACKPTVCSMP SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QQESPAAGAARMNCKEGTDSSCGCRGNDEK
						KMLKCVVVGDGAVGKTCLLMSYANDAFPEE YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH RRDQKWHDKQYKKAHLGTALKANPFGGAS HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK NGKKITAFVPNDGCLNFIEENDEVLVAGFGR KGHAVGDIPGVRFKVVKVANVSLLALYKGK KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME TQSEPSELELDDVVITNPHIEAILENEDWIEDA SGLMSHCIAILKICHTLTEKLVAMTMGSGAK MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL DPKLLDARTTALLLSVSHLVLVTRNACHLTG GLDWIDQSLSAAEEHLEVLREAALASEPDKG LPGPEGFLQEQSAI
1059	2409	A	8809	246		MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTVHNPGLALLHLLLLYGLVVSTALI WHPINKLAALLLLPYLAWLTVTSALTYHLWR DSLCPVHQPQPTEKSD
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA FIFSYITAVILHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATIYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824	1	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNITSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV
1063	2413	A.	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR
1064	2414	A	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	,	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ĺ	ì		amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
1	1			sequence		nucleotide insertion
<u> </u>	 	<u> </u>		Sequence		VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
		1				SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
		ļ.				EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
	ŀ					QAQEELAEMKRYSESSSKLEEDKDKKINEMS
	ļ .		ļ]	KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
]	}			LQQQVKQLQNQLAECKKQHQEVISVYRMHL
1]					LYAVQGQMDEDVQKVLKQILTMCKNQSQK
ł						K
1065	2415	A	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
1005	2413	^	0041	,	003	APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL
						KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
	1	ł i	Ì			TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
						DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1	İ					DPOTLKPSGFHEDDPFFYDEHILRKRGLLVA
				,		AVLFITGIIILTSGKCRQLSRLCRNHCR
1066	2416	A	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG
1000	2410	^	6633	3800	2204	RRRRGRVVSRKKMSLKSERRGIHVDQSDLL
1	1	l				CKKGCGYYGNPAWQGFCSKCWREEYHKAR
						QKQIQEDWELAERLQREEEEAFASSQSSQGA
						OSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
	İ					VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE
		!				FLKTFHKTGOEIYKOTKLFLEGMHYKRDLSIE
						EQSECAQDFYHNVAERMQTRGKVPPERVEKI
						MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
		1				QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV
1						KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
						KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
1						QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
						KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
[[WSPDACLGVKQMYKNLDLLSQLNERQERIM
		1				NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
						KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	Α	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
						LRQAWATKQDPISKKK
1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	Ā	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	A	8866	293	1675	PYPQGGYPQGPYPQEGYPQGPYPQGGYPQGP
					1	YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
	Ì					EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
						LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
1						WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
						VALSVLTASLSYMVGMIASFYNTEAVIMAVG
						ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
						VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
						VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI
						FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
						WHGSASCTSPLSCPQAQPREKDASLQPSCMY
						TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
1						HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
						EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
						GDMRSGGLIPVLSPE
1071	2421	Α	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
						DDKMGSNTFFKRNDCRYVMISCKADMAYDN
1						VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII
		N				LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	Α	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
	- · 					GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI
1						GGTGPSSDAGWGCMLRCGQMMLAQALICRH
						LGRDWSWEKQKEQPKEYQRILQCFLDRKDC
		لـــــا				

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq-	seq- uence		USSN 09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	uaice		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1 40,100		l	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł	!	peptide		/=possible nucleotide deletion, \=possible
		1	<u>i</u>	sequence)	nucleotide insertion
						CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ
		l				GV*KNLA/LFDEW\NSLGLVYVSM\DNPSGSIA
1050	0.400					RFPKKLCRVLPL\SADTAGLTGP
1073	2423	Α	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
				:		*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
1.07	2721	1.	0004	0,	133	KEISFGDYICHTFQGDCWADRSPLHEAAAHG
1						RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
						*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	Α	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
						SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
1			ľ			PWPSLLDKEREESLRQKRLSERERIGELGAPE
						VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
						TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
1						QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
						QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
						PRRGEIGLTR*RNCHHLNAQVM**VVSRIIRR
					İ	MEAVRTAKREPESTVLMRREPLHPFNPRRET
						KERE
1076	2426	Α	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
						*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
						FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
1 1						VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR
1						VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
1 1						RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
1 1						QYPALHRAGTEWOLSALHRAPRSTOPDKAC
l						RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
						\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	Α	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
						PALPFAATPGSRGQALCRGGRRRQHLHGPLH
1079	2429	Α	8912	121	376 ·	RP*QAAPALHAGCQLAPHPPT
10/9	2427	A	0712	121	3/0 .	NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
						EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
						GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
						DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
 	İ				İ	YAPICMEYGRVTLPCRRLCQRAYSECSKLME
				{	ĺ	MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
					ļ	GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
	1					SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
						IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
			-		•	WHMMVSLIFF\GFLLEDRVACNA\SIPAQYKA
						STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA
	ľ	1	· I			WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
				ł		VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
	1	ł	1	ļ	Į	VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
						VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	Α	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
						CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
	1	- 1	I	ľ		TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1082	2422		9022	255	1070	LSIAIPAMVNNTAPPSQPNASTERPST
1002	2432	Α	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	-	/=possible nucleotide deletion, \=possible
		1		sequence	1	nucleotide insertion
						GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
						QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
		l	ł	ł	ł	FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
				1		EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK
	•					HNNQILKTSYA*HQQS/RQIQKKMMEIMT*EV
		1		1		QTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH
						DVFIRKVKMLENPGFER\MELRGGGSSS
1083	2433	Α	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
				i		WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
	1		ł		1	AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
						GWLLT/ESCSISPKLCSIAVH*DNPAWF
1084	2434	A	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
						WGGKRVQPFWKRVWQKRTLNLRV
1085	2435	Α	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
		ļ				*TIYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR
						MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
						HKYYIKVKRNL*VTITH\TWVNLNILMFEIILW
1004	2426		0000		1006	YSHKYY
1086	2436	Α	8962	868	1026	H*KILQVGRAQRAHXSRL*SQLLRRLRHESHL
1007	3437	 	0005	50	220	NPGARGCSEARLHRCTPAWTT
1087	2437	Α	8985	58	330	LHVKHLGHFQLVFSEVICHCILMPVS*ELQRL
	l	l		}		*ERSVCAFHVCIQTYVCLQVYACMCVYYICM FVYSVYGCGLCTCVCMDVYICVCVQEFL
1088	2438	A	8989	394	404	N*KWILHVNVRIQSIFF/IKRNQK/INSHELKLD
1000	2430	A	0909	394	404	KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA
	1	Ì				KYTVKRIKIIHPTDLEKMLRNHLSDKD*YS/GV
					İ	YKDLSKLNRRKTE/S*/VKKWVKDLSRYFIKE
	i	ĺ	l	Ì	i	VISMENKHKKIFSTS
1089	2439	A	8991	60	329	MALTPESPSSFPGLAATGSSVPEPPGGPNATL
	1		"	1 00	527	NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
		1	l	ļ		GVEDNAYTLEVNSRYMRAVGIM*IHL
1090	2440	A	8996	2	351	SNITITLT+MKKYDNTFCW+GCGQIG/T/LIYC
		••				WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL
		ļ		ļ		GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR
		l				LLIAALFIIVQYWKQSKDHYI
1091	2441	A	8997	97	456	YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT
		1				LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG
						AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS
	1	ļ	1	·		FDKNYKAPIGADFEMERFEVLGIPF
1092	2442	Α	8999	548	811	SSFIKRHILIFEDDWHQTTCCHHPHHP\F*RCQ
		1				FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH
		L				RAAHHHQHGQGPLGHGLVARVG
1093	2443	A	9002	3	2745	ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS
	}	1				TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD
	1	1	1		[AMCCLRYWYTPESWICGGQWREYFSALRDF
		1				VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR
	ł	ł	}	ł	}	LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA
	1		1			VFTRFALKTLGQETLCSLQEADYEVASYGLQ
		1	ł		1	HNCLGILGEDTDYLIYDTCPYFSISELCLESLD
	l	i	l	[TVMLCREKLCESLGLCVADLPLLACLLGNDII
		1				PEGMFESFRYKCLSSYTSVKENFDKKGNIILA
	ļ					VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL
	1]			*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP
		!	İ			RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP
· 	ł	l			i	GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG
			1			PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP
		1				MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP
		1				RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY
	L	L	l		l	TGPESRQEVLIRTDPESRQEIMCTGHESKQEV

No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide			1	'			
194	nucl-	peptide		in	nucleotide	location	
194	eotide	seq-	į.	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
mimo acid residue of peptide sequence	seq-	uence	1	09/496			
Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide Sequence Peptide	uence			914	ng to first	acid residue	
Populita							
	1	}	ļ	ļ		sequence	
PICTOPISK OBSMCTHAEINOKLEPVATIDEER	,	į.	1				/=possible nucleotide deletion, \=possible
LEALMCTNPERQGEPTNNGGEVGQOVTMNS DTEILK/NARTHHVQAESHLYNNINSGEIECS			<u>. </u>		sequence		
DTELLKVARTHHVQAESYLVTYMIMSGGEIGEN NTLEDELDQALPORPROSAPITYRPIGORYSILLED							
	}	1	j	ļ	ļ	İ	
CODVTSTCLAVKEWFVYPGNPIRHPDLYRG CODVTSTCLAVKEWFVYPGNPIRHPDLYRG CONTROCTPSIKILWINOEPEIOVRRLDTILA CFALSSSREILQAVESPPQALCCLLIVILFYD DTLCLEDLIHAFIAQALCLQGKSTSQLVNLQP DTLCLEDLIHAFIAQALCLQGKSTSQLVNLQP DTLCLEDLIHAFIAQALCLQGKSTSQLVNLQP DTVTNPRAVQLOSLLYRGLTITLVLNSACGFP WKTSDFMPMWFDGKLHFQKNLQSEKGYQ VEVL/CRK*ISAGIQPQEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV VEVL/CRK*ISAGIQPQEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV THMAGEVSPLKHFVLAKKAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHFVLAKACAITAFPQLIEAT THMAGEVSPLKHF							
			}		ļ	İ	
		1	})	!	<u> </u>	1
DTILCLEDLIAFIAQALCIQGSLLYRGITULVLVNACGP DYNPRAVQLGSLLYRGITULVLVNACGP WKTSDFMPWNVPDGKLFHQKYLQSEKGYA VEYL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGITPRREVGKTGLQLPQDGLWV VEYL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGITPRREVGKTGLQLPQDGLWV VEYL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGITPRREVGKTGLQLPQDGLWV VEYL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGITPRREVGKTGLQLPQDGLWV VEYL/STREMKVAVFGGTISAHCALAHQMCDLK VEYL/STREMKVAVFGGTISAHCALAHQMCDLK AREACRAKTDFPGRRFRLWPSCCCTVGAE THAMAEVSLKSVKTVTNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD VTTELDSWIDKFCTKSSTREITNSGGSSV NAM. WDKVLPSGIGHITNCFLSVEGTIGDKA YLMSTREDFPPEGAGRITPFALRPLAACW LHRRARRSSALCPRPSWOVSGGEGAGARE P*TSSSCCLSAA/SHLSIQSPMAGARRRIRRQ LAKEKIEGCHICTSVTPGEPQVFLGKDKATF DVVTDIDSQCQFOVINCIERLIGECFGYNATV FAYQGTGAGKTYTMOTGFD AREACRA SHENGSPMAGARRRIRRQ LAKEKIEGCHICTSVTPGEPQVFLGKDKATF DVVTDIDSQCQFOVINCIERLIGECFGYNATV FAYGGTGAGKTYTMOTGFD TO VTTELDSWIDKSCFTSKLKKLAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRE*SCPTTAWTTERDPV TO VTTELDSWIDKSCFTSKLKKLAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRE*SCPTTAWTTERDPV TO VTTELDSWIDKSCFTSKLSLASPRVKCSGAI LAHCNFRHAGFPLSCLSLPNRWCSGTGALDLTSRSACLG LPKCWDVRREPAASIIGTTFFINSK CONTROL TO VTTELDSWIDKSCFTSKLKVLAFLCKRTS TNPSQGPYHLWPSHIFWQTTCCRLPHTKQ G*AALDHLKVPSRIPLPJVKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGKRKVYTQKLCYQKK TSSWDYRRGWWGGGDIETLHCW*E*KII TSSWQCGTSMLNIAHHNQK WKKINI THE STREMGCMAW SATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGKRKVYTQKLCYQKK TSSWDYRRGWWGGGDIETLHCW*E*KII TSSWDYRRGWWGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRWWGGGGTSMLINAHDQKWKKNII TSSWDYRRGWWGGGGDIETLHCW*E*KII TLJFFSMQCAKRACHAETLILLE TLJFFSMQCAKRACHAETLILLE TLJFFSMQCAKRACHAETLILLE TLJFFSMQCAKRACHAETLILLE TSSWDYRRGWGGFOSTSMLNAHDALNGLWKKII TSSWDYRGGFOSTSMLNAHDALNGLWKATILLILLE TLJFFSMQCAKRACHAETLILLE TSSWDYRGGFOSTSMLNA		1				Į.	
DYNPRAVQLGSLLYRGITTLYL.VNSACGP WKTSDFMPWIPDGKLHQKYLQSEKGYA VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV TEGSHFVBATTKNPELDBILYBSCCCRVIVGGAE TPIMAMEPVSPLKHFVLAKKAITAITPOLLERV TEGSHFVBATTKNPELDBILYBMQGYV KDKLSIIGEVLSRRHMKVAFFGRTSSGKSVI NAML WDKVLPGGIGHTDFCLSVEGTDGDKA YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALRDDLVLVDOPCTD VTTELDSWIDKFCTKSSTREITNSGSDT LVLNSRVEDFVPPEGAGRTLPFALRPLAACW LLHRARRSSALCPRPSRVGGEGGAGAKE PPITSSSCCLSAASHLSIQSPNMAGARRIRPQ LAKEKIEGGHICTSVTPGEPGYPLGKDKAFTF DYVPDIDSQQEQIVIQCIEKLIEGCFEGYNATV FAYQQTIGAGKTYTMGTGPD TEGSTOP TEGS	i	j		İ			
WKTSDFMPWNVFDGKLFHQKYLQSEKGYA		1	l		ł		
VEVL/CRTK*ISAHQIPQFEGSRLQGLHEGGOT			1				
HHWPSPLGLTPREEVGKTGLQLPQDGLWV	Ì	1		!		}	
1094			i	i	Í		
TEGSHFVEATYKNPELDRIATEDDLVEMOGY KDKLSIIGEVLSRRHMKVAFFGRTSSCKSSVI NAMLWDKVLPSGIGEVLSRRHMKVAFFGRTSSCKSSVI NAMLWDKVLPSGIGHTNCFLSVEGTDCDKA YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRYPWFKAKCALRDDLVLVDGFOTD VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT VLNSRVEDFVPPEGAGFTLFFALRLACW LLHRARRSSALCPRPRSWGVSGEGAGARE PPITSSSCCLSAA/SHLSQSPNMAGARRIRPQ LAKEKIEGGHICTSVTTGEPQVPLGKDKAFTF DYVFDIDSQQEQIVIQCIEKLIEGGFEGYNATV FAYGQTIGAGKTYTMGTGPD FFFFNVCKSFKVPRFGCKEESTGTLFKNTLISL GQHSETPSLKKKLAGYSGMCL*SQVLRILRQ EDCLSPGGONCES*SCPYTPAWITEDDPV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCPYTPAWITEDDV EDCLSPGGONCES*SCP	1094	2444	Α	9021	97	834	
			1		'		_ <
NAMI_WDKVLPSGIGHITNCFLSVEGTDGDKA				•			•
1095	ŀ	1]	
AGCLVRYPWKAKCALLRDDLVLVDGPGTD		ļ					
VITIELDSWIDKFCTKSSTREITNSGSDT			ļ				
1095							
LLHRARRSSALCPRPRSWGVSGGEGAGARE	1005	2445	<u> </u>	0000	<u> </u>		
P*ITSSSCCL\$AA/SHL\$IQ\$PMMAGARRRIRPQ	1095	2445	A	9022	4	337	
LAKEKIEGCHICTSYTPGEPQVFLGKDKAFTF							
DYVFDIDSQQEQIYIQCIEKLİEGCFEGYNATV		ļ	l				P*11555CCLSAA/SHLSIQSPNMAGARRRIKPQ
1096		İ	1				
1096		i	İ				
GQHSETPSLKKKU_AGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRE\$*SCPYTPAWITERDPV	1006	2446		0020	1	205	
EDCLSPGGGNCRES*SCPYTPAWITERDPV	1090	2440	^	9029	1	203	
1097							
LAHCNFRHAGFPPLSCLSLPNRWEYRRPARP GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASIIFQTTFFINSK 1098	1007	2//7	Α-	0032	716	357	
GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG	1077	2777	Λ	7032	710	337	
LPKCWDYRREPAASIIFQTTFFINSK 1098 2448 A 9038 230 652 KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRVIWQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLLFEMESLPVA/RVECSGTISAHCNLCPLGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDSLLMKLYSFLLNDSPLNPLLASFF SKVLSILJSRKPFQQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSIL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC/GA			ļ				
1098							
TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ	1098	2448	A	9038	230	652	
G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF R*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA	10,0	1	••	3030	250	032	_
VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVPTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA							
TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFIPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLLFEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA							
1009 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQK WKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLLFEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA							, ,
HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLLFEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS HID DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILJSRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA	1099	2449	A	9043	185	372	
RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVOFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISK YVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA							
1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLL FLLFEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVOFLKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA							
SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLL FLLFEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVOFLKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA	1100	2450	Α	9045	763	584	
1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF I\(\delta\) EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVOFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA			-/.		,		
FLL/FEMESLPVARVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCAGA	1101	2451	Α	9050	275	2	
DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA							
1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF I*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA							DSPASAS*VAGITDMCRYTQLILFHAS
DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF N*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA	1102	2452	Α	9053	449	1224	
IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLRLLTCIEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA							DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF
NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA							
SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA]					
RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA				{	ſ		
RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GC\GA		[Ì		ļ	
1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA					l		
1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA						ļ [LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS
WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA					' [SLV
	1103	2453	Α	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
AGMOI/H/CW\WCVNVGKFWEMS*YYLLKLSI					ļ		
				ļ			AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI
ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA							ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
иепсе			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		/=possible nucleotide deletion, \=possible
				sequence .		nucleotide insertion
						APFVLAVNC
1104	2454	Α	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI
i						KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V
		1				KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
						YMQIGTTKNSRAT
1105	2455	Α	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS
1						RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR
l						AWPCCPGWSAAWLTIVILAHYRRPGLERSCC
						LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL
	2155		2000			VLNS*TQGI
1106	2456	Α	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT
1100	0.455		0001	600		HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT*
						AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA
						VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL
	i					I/HCWWEYNVIHIWNSVTFPRKVEHVYITYA
J						PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	Α	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP
1100	2436	A	9093	340	1	GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA
1						SAFPPAERSRGHRRASL*RARWSAAVPRRSA
Į.						,
ļ						GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG
						QRPPPPSGDSLSPPGCCRY
1109	2459	Α	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL
1105	2433	Λ	3033	1233	1423	GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	Α	9103	242	70	EEOFFFFAVGMFP*VDFLAPASGELWDRLRLT
	2100	11	7103	242	/*	CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP
1						SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS
						LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP
			,,,,,	-00	,,,,	AAAGDPASLDFAQCLGYYGYSKFGNNNNYM
						NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT
ľ						PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT
i i						POFPPOSLDLPSITISRNLVEQDGVLHSSGLHM
						DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG
						VMPPAQLTTINQSQLSAQLGLNLGGASMPHT
						SPSPPASKSATPSPSSSINEEDADEANRAIGEK
						RAAPDSGKKPKTPKK
1113	2463	Α	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F
						SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH
						VGQAGHEPLTSGDPPASASQSAGITGVSHQA
						WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS
						LLKCWDY
1114	2464	Α	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS
						SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII
<u> </u>						YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG
						SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM
						TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE
		1			1	WRLQHKGRGRGDLHLPDHHLSVPSSADHPA
				_		QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA
		1				QHVHVPPWTDVLAGQDRRAPTAGDGAPWP
[APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ
						PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA
						GAGGP*GSPAGRACGAAGCRPRPPRPAASSA
						*NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTITLWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*GILTKLSAFLTIPRLQPHLIAALSPSS
1118	2468	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	2		ACPRLARRRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSVLRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQGURG\ WKARRAITTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEEDEDDEGGEEAPAPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1121	2470 2471	A	9163 9166	124 272	523	PPRACRPCPRACPCPPT*KCSQPVSWPC PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
						V/CSHITDSLKFIGKGWVGMVTHACNPGTLG
1122	2472	С	9170	442	236	G*GGWIA*VREFETSLGNM MNRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	,	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-]	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
1124	2474	A	9173	sequence 3	374	nucleotide insertion GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
1124	2474	^) 1/3	'	374	WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
				-		TGSLFVLLGVFSFEPVPSCRALOELKPRDRISA
ľ						IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
Ì		ļ				CSKPPKETGELENAESGGDGGRRGGKQDNV
ĺ	İ					AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
		1	l			LPMGFFYLYFRDPGREITWKHFVQYYLARGL
ĺ		[ĺ		ĺ	VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	A	9185	1	321	LTGOLGSILLRVFSKSRAGLGARKLKAYRTM
				-		EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
	1	į			·	ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
						LTFFSTIS
1128	2478	Α	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
						RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
Į,	i					LVLDGVPVALKKVQIFDLMDAKARADCIKEID
						LLKQLNHPNVIKYYASFIEDNELNIVLELADA GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
						ALEHMHSRRVMHRDIKPANVFITATGVVKLG
		ļ				DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
						NG
1129	2479	Α	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
						PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
		1				RATTIKIRVVATITRARIEDMRHSATALTRPD
1130	2480	A	9194	131	487	ATTAQIPKLPVTTVCNRRANPGIPPSVL AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
1130	2460	А	7174	151	407	LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
						PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
						DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
						PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
						CVNKTESNLSIDIAFAYPFRLHQVIFEG\PTCE
						GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL AATGVYIFFONKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK
1	2402	1.	7200	•		TQPVEATDDAFWDQFWADTATSVQDVFALV
						PAAEIRAVREESPSNLATLCYKAVEKLVQGA
						ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
						WRGFFWSTVPGAGRGGQGEEDDEHARPLAE
						SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
						SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME LLKLLLTCFSEAMYLPPAPESWOH/RTHWFSS
						FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
						NHLY
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
]]						AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
						HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS
- <u></u>						NVYFIV
1134	2484	Α	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
[]	ľ					RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
						AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC GVFSIVGALCYAELGTTISKSGGDYAYMLDV
						YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
	}					LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
						YSVKAATRVQDAFAAAKLLALALIILLGFVQI
						GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
Ll						LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII GMIWLRHRKPELERPIKVNLALPVFFILACLF LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ RDRLPPAYFCRPVVCVVTALDVG\SPESQEM
						DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKNPRNNL RSLLGERVDENTEENHCGETSSQIPDDTLNK
1136	2486	A	9223		983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG DLVFAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD SEAPEANPADGSDADEDDEG\RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK MSVSKRARKASSDLDQASVSPSEEENSESSSE SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK PKPERPPSSSSSD
1137	2487	A	9229		239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA
1138	2488	A	9231	207	443	TRSVGVNICEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR ARSAAAAPGSEASFTESRGLQNKNRGGANGK GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TIPQGKPETIFLDQGCSSPVLIDCPHPNCNKK YKHINGLRYHQAHAHLDPENKLEFEPDSEDK ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK ANNCKTDKNPSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
						VTAAVSGLLVGYELGIISGALLQIKTLLALSC HEQEMGVSSLVIGALL
1140	2490	Α	9238	248	328	MAQGNNYGQTSNGVADESPNMLVYRKV
1141	2491	A	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA ATWGRLPGPEETLPGQDSWNGVPSRAGLGM WPWAAALVVHCYSKSPSNKDAALLEAARAQ WMQEVSRNRCALLHSAAVQEYGYGN
1142	2492	A	9245	157	466	HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF GICKEYSRQ
1143	2493	A	9247	264	115	GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG ARDSTSIIRMGPEIPPP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ĺ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Í	[ĺ	ĺ	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1144	2494	Α	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP
ļ	1	l				SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE
		ļ			İ	AMEESDRPCEISEIDDNPKISENPRRSPTHEKN
			Į.			TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM
						ERRR
1145	2495	Α	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF
i	İ	İ	İ			PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL
		•				WDTAGQERFISIT
1146	2496	A	9277	592	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI
						SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA
						WLIFAFLVETGF
1147	2497	A	9279	1255	2	FRRGRRGEEEKEEEEEEEGWVNGMENSHPP
1			1		ļ	HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP
Ì						ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE
						DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV
						PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE
1			ŀ			VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ
ĺ	{	ĺ	i			HLHSTSVMGNIIHVELDTKGETRMRFYEL\LV
						TGRYTPQTLPVGELDAVSPIVNETLQLSDALK
	ľ					RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA
j]			YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY
			ļ			SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY
ŀ						NRRHEHHYVHNSPAVTAVAGATAAFRGSSD
						LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP
1140	2400		0202	1006		PAQATPAPGFR
1148	2498	Α	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY
						ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL
						SMCLVT:VLGNLLIILAISPDSHLHTPMYFFFSN LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG
						CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI
1						CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL
						HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD
						SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL
					,	GNLIILDVSPDSHLPTPMYFFLSNLSLPDIGFT
						STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF
						GGMEERHAPECDGL
1149	2499	A	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE
			7505	•		FRLVAADRSMGRYMLFGVINLICTGFLLMWC
						SSINSIALI\SYTYLTIFDLFSLMTCLISYWVTL
						RKPSPVYSFGFERLEVLAVFASTVLAQLGALF
1						ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF
						TMLSIRNKPFAYVSEAASTSWLOEHVADLSR
						SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT
						YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL
1				·		SLP
1151	2501	Α	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPOTOORAG
						SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ
						FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV
						QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR
						PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I
						PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL
						GTNVSLRAA
1153	2503	Α	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP
				-		PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG
						GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA
						PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR
					ľ	PGNS
						· · · · · · · · · · · · · · · · · · ·

OFO ID	CECTO	Mat	LCCO	Dendies -	D 3: -3 - 3 - 3	I Amin and a series of the control o
SEQ ID	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	peptide	1,100	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ï	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l	l	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
			1	sequence	1	nucleotide insertion
1154	2504	Α	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
			<u> </u>			PT
1155	2505	Α	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
	ŀ	ļ	ł		}	EAAQRHCRASVSILRMRRPGQGSSRPARVPL
		L			222	RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGGSHPAGAQPGGRAIGAGWQS
1150	0.700	<u> </u>	0000		100	KEPLWEGLQRSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
						LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
						VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
	i	i	1		}	RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
1133	2.503	^	7554	100	363	LMEAGLPQKQAERADELFEAGLVIYVKLDER
			ŀ			VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
		1	1000	_		KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
			-			DHTDQELREEIHKANVERVVHDVSQEATIEKI
		ł	ļ			RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
						EAELPIMSQLTEIETCVEC
1161	2511	Α	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
						AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
			ŀ			DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
	İ	1				QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
1162	2512		9343	84	837	GELI
1102	2312	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
					•	VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
	}	1			}	VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
		l				GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
	1				,	FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
		ĺ				WIDGPKDTSVEDALERTYVPCLKTLQEEVME
		[AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
						G
1163	2513	Α	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
	1	l	1			LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
		1	1			AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
1164	0514		02.45		1000	FSSYQRNNPDLILNDTIMPNIK
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
		l [']]			HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL FSVLLPLRLDGIIOWSYWAVFAPIWLWKLLV
		1				VAGASVGAGVWARNPRYRTEGEACVEFKA
		l				MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
]	ł				VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
		l]			SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
		l				SFLCLVVLYYIVWSLLFLRSLDVVAEORRTH
	}]]			VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
		1				YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
		Ī				AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
		_				LPLQNKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
		I				VPEGVRLADGPGHCKGRVEVKHQNQWYTV
		l				CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
						TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
1166	2614	<u> </u>	0262	201	200	LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFFESEFY/SSPRVECS
1167	2517	 	0269	707	1007	GAISAHLAHCNLCLPGSSDSPASAFQVAS
110/	2517	Α	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

OPO ID	L CEC ID	1 1/	I CEA	Dungling	Daniel Salara C	LA-in-orid
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	,		į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ	į	ŀ	l	peptide		/=possible nucleotide deletion, \=possible
			ļ	sequence	Í	nucleotide insertion
						PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
		l		ł	[ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
						PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS
j			ļ	,	ļ	PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID
		•	}			LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS
ł	1	1				QQSILAGLVVVATTGMIGSPLECLFGELGGRA
	}					DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL
	L		<u> </u>			FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN
	}	1				NSARRMEAMASGSNWLSGVNVVLVMAYWS
		1		1		LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH
		ļ.,. <u>.</u>				NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170	2520	Α	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR
ĺ		ĺ	l			ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF
						TNETWQARTGEPLPDHLVLLMWSLIVSLYPL
		l				GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS
		l				AAILFGFSRKAGSFEMIMLGRLASWGVNAGV
		İ	1			SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA LGIVMGQVVGLSTTAATGLRGL\AGELEELEE
		1				ERAACQGCRARRPWELFQHRALRRQVTSLV
	ł	ł	1			VLGSAMELCGNDSVYAYASSVFRKAGVPEA
						KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL
						WGGTPRSFALNQFTLQKKKK
1171	2521	Α	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA
****		' '	, , , ,			LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE
			1			TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV
						EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG
						SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKET\KAAMGTQ
						CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL
						LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL
						ASFNEVGNTALIVLESY
1173	2523	Α	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ
						QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI*
]	KIFVLFDFNIMFETPFYII+FIFLFSQNLKRIRQV
						IRPPISFSKINNGP
1174	2524	Α	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ
						RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF
						ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF
1100	2525		0700		207	LGML
1175	2525	Α	9399	66	397	HESSRADROKMDTRGSTYTDADPVNKSGGT
						AKMNKWSKGKVRDKLNNLVLFDTATYDKL
			ļ	'		CKEVPNYKLITLAVVSERLKIPGSLARAALIIE
1176	2526		0409	<u> </u>	200	LLSRGLI*LVIQHIAQVIY
1176	2526	Α	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY
						GERGYAQNGDF*DAQLDDYSFSCYSHAQVN
						GAPNSLTRAYDDP*VKISGLECQKVGALVEV
1177	2527	A	9416	2	402	KCLNL CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV
11//	اعدد	Α	7410	-	402	YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD
						FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF
1178	2528	A	9419	142	426	DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1110	4346	A	7417	146	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL
				ļ		ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG
,	2327	^)72V	1750	1000	GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA
1						WWRAP
			لــــــا			

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN 09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence		Ī	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1		peptide	l	/=possible nucleotide deletion, \=possible
		[.	1	sequence		nucleotide insertion
1180	2530	Α	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
		l				SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
						L
1181	2531	A	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
						IARTILDRLTGIPHGYCFVE*ADWATADKCVH
						IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
1			1			SMILK*MGAGDEKISAMGKARVDHRELYLGL
1102		<u> </u>	0444	204	2	LYPTEDYKLTFRARH
1183	2533	Α	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR
1						KGCSGWAPWLSLQCQHFGRPRWADHLRSGV RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
	1		}			ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
	l	l]			ERG
1184	2534	Ā	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG
1		l '`	1			RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
						IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	Α	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
Ī	•	l				WWWGWECWVRALKLSSGPAGPLACWVAK
		1			! !	KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
	l	<u> </u>				WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	A	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR
						GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	Α	9469	388	3	EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ
		1				SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
1	ŀ	İ	-		1	NPASPHPEAPQEPWDSASGSVGSFSLGRGAK
]	J			ASS*VPGKGRGPRQGSELLAETILELFLALAN S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
1	2230	l	´ ′′′			GRLMANPEALKILSAITQPMVEEAIAGLYRAC
			<u> </u>			*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	Α	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
	ł	l				PSLLKIQKISWAWWRAPVVPATWEAEAEEW
		l				R
1190	2540	Α	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
						PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR
	1					GASSCRRRCNPVLAARKAGSPRSHSTRENC
110:	2541		0400	<u> </u>		RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	Α	9489	I	411	LADALCLSAAATGAVRPGARAQPSTRRLSP
						SVRVCCRAAAASNLLYSSCLQRHSERASEEG ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
		ŀ				MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ
		[KEEELTAVNVK
1192	2542	Ā	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
	<u>_</u>					CEEDERKMAREFLAEFMSTYVMMNIHMIVE
					ļ	KDTYSDHEEINTS
1193	2543	Α	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
						FGKTFVNVN+S+TYVYPCDKIILLLGLYPTEM
1194	2544	Α	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL
		l				LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
						SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
						RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	Α	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
						AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC
						PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
		l				LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
						PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP APGSGPCGATARPSRGGRAGGSRARRPIPPGP
		1				GTRRTPSGCONPAASGG
L	L	L	L	L	L	o i i i i i i i i i i i i i i i i i i i

SEQ ID NO: of NO: of NO: of nucl- NO: of NO: of nucl- nucl- peptide eotide seq- uence USSN location or sequence (A=Alanin nucleotide location peptide eotide seq- uence USSN location or sequence (A=Alanin nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide location corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide incation corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide incation corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide incation corresponding to last amino acid residue of peptide sequence (A=Alanin nucleotide incation in nucleotide incation nucleotide in serior (A=Alanin nucleotide incation to last amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation in nucleotide incation nucleotide in serior (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide incation (Amino acid residue of peptide sequence (A=Alanin nucleotide insertion (Amino acid	acid, =Histidine, cine, P=Proline, erine, rytophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE PNKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
nucleotide eotide sequence peptide sequence in undestide location corresponding to last amino acid residue of peptide sequence Image: sequence sequence sequence In undestide location corresponding to last amino acid residue of peptide sequence F=Phenylalanine, G=Glycine, H=Isoleucine, K=Lysine, L=Leuc M=Methionine, N=Asparagine, I M=Methionine, N=	=Histidine, sine, P=Proline, erine, rptophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE *NKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
eotide sequence	eine, P=Proline, erine, erine, rptophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE PNKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
sequence uen	P=Proline, erine, erine, eptophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE *NKRHRVAEWI *RLKEREQKKRK FIGPPKSIPWAA GTX
uence 914	erine, ptophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE NKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA GTX
amino acid residue of peptide sequence y=Tyrosine, X=Unknown, *=Str y=Tyrosine, X=Unknown, *=Str y=Tyrosine, X=Unknown, *=Str y=Str y=Str y=Tyrosine, X=Unknown, *=Str y	rptophan, op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE NKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA GTX
residue of peptide sequence	op codon, =possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE NKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA GTX
peptide sequence /=possible nucleotide deletion, \= nucleotide insertion 1196	-possible GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE ONKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
	GGRPLPLLGAM SRRTPHL**EPHA YKLTADSRYRG OSRFQVQRARYE ONKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA
1196	SRRTPHL**EPHA YKLTADSRYRG SSRFQVQRARYE PNKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
AERVAPGWDLHTPYLPRTNS GYIGALFPMSGGWPGGQ	SRRTPHL**EPHA YKLTADSRYRG SSRFQVQRARYE PNKRHRVAEWI TRLKEREQKKRK FIGPPKSIPWAA GTX
GYIGALFPMSGGWPGGQ	YKLTADSRYRG OSRFQVQRARYE *NKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA GTX
1197 2547 A 9521 289 448	PNKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA
HAMRHLTGNTSMAIRFL*AD	PNKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA
APNWKYKYGY*IPVDMLC 1198	*NKRHRVAEWI RLKEREQKKRK FIGPPKSIPWAA GTX
1198 2548 A 9524 204 I KNKKTTKCLSIVTLNISGPNQ	RLKEREQKKRK FIGPPKSIPWAA GTX
VKQEPNICHL*ETHFPFRDTY SSYS 1199 2549 A 9546 1785 1943 GGRFKESKLTNAGWQRNSFI V*QRGDGKNPGVTHLNRPVC 1200 2550 A 9548 186 1 VNAEKEF*KIQHYFMTKSQN KAIYDKWTSDIMLNLQKL*A	RLKEREQKKRK FIGPPKSIPWAA GTX
SSYS	FIGPPKSIPWAA GTX
1199 2549 A 9546 1785 1943 GGRFKESKLTNAGWQRNSFI V*QRGDGKNPGVTHLNRPVC 1200 2550 A 9548 186 1 VNAEKEF*KIQHYFMTKSQN KAIYDKWTSDIMLNLQKL*A	GTX
V*QRGDGKNPGVTHLNRPVC 1200 2550 A 9548 186 1 VNAEKEF*KIQHYFMTKSQN KAIYDKWTSDIMLNLQKL*A	GTX
1200 2550 A 9548 186 1 VNAEKEF*KIQHYFMTKSQN KAIYDKWTSDIMLNLQKL*A	
KAIYDKWTSDIMLNLQKL*A	TAY BETTO PRINTED TO TAKE
1001 0751 1 0740 501	
1201 2551 A 9549 591 2 SSVVEFPRGPRSSLPPLDSTFF	CGSSPNWTGGC
GSCPSGE*LVSPGSEQRKKYS	
YHVQHLATFIMDKSEAITSVI YHVQHLATFIMDKSEAITSVI	DDAIRKLVQLSS
KEKIWTQEMLLQVNDQSLRI	LDIESQEELEDF
PŁPTVQRSQTVLNQLRYPSVI	
PDVHFFHCDEVEAELVHEYN	
AMRP	
1202 2552 A 9552 428 1 KYGNEGHWSRQCPNPGKPIR	PCPLCRGPHWK
LDCERPPOGPLPSLPELAKTS	
*WGPGMDAPATTIASSKTRV	
LI*YRATYSALPNFSGPTQSSG	OVSVVGIDGOV
SKPRATPPLFCSLHTF	` ` `
1203 2553 A 9568 517 738 RRKFERKQKQ*RYREGKQYF	RORDKMKEWG
EKEKRREKGEREERKMRHI	
TMENWRVERLTEKER	
1204 2554 A 9573 83 415 EDKRLRLVDGDSRCAGRV*I	YHDGFWGTICD
DGWDLSDAHVVCQKLGCGV	
EGSGPIWLDDLNCTGTESHL	
HDCRHKEDAGVICSEFTALR	
1205 2555 A 9577 64 424 ARGSCPTRPRTANGRMGETK	
DVAVTFFREEWRQLVLVHR	
GLLDTLRHNVPQPDVVHLLY	
VSHSPCAGDMRELFTREATL	
1206 2556 A 9584 38 476 TLGAVLFSEVSKESSTSHSGG	
SNFITPSSPRLKP*TASSQRNL	
NPQPLSTPSWQIETKYSTKVL	TGNWMEERRK
GLPYKHLITHHQEPPHRYLIS	
YNPGLPPLRTWNGQKLLWL	
1207 2557 A 9586 2 412 LRSSPAALLRALCITTVTGTA	LALRSRVATTN
PDGCRNVLRPKYYRLCDKAL	
GVAVTSWAIMLTVLTLVCKO	
THILCLL*EKGIFGLTFAFIIGL	•
FGILFSICFS	mooi di imili
1208 2558 A 9597 122 3 IKNYWPGMVAHACNPSPLGO	SDCDWA * AOV
1208 2558 A 9597 122 3 INNI WYGMVAHACNPSPLOC	AVW.WIM.OVE
	ECHCCDTCTDY
GDPGPTFSKMSIWTSGRTSSS	
RIRDHDLLDKRKTVTALKAG	
MVCSIMM*FLLGITLLRSYMO	ASAMLKEZÓCL
LLNASITETFNC	
1210 2560 A 9618 384 2 SLHDMLMLAEQQQKQKWAY	
DSKFGQRILEKMEWSKGRGL	
KVQVKNNDLGLQATINNEAN	
LLAELNTCQRQETADS***WS	3PKNSHVGKDS
GELSAK	
1211 2561 A 9620 316 610 QKHPGGGQLGRSPQEDSRFH	NKASSGVSRVR

OTO IN I	020 70	137.	1 000	(N . 1'		· · · · · · · · · · · · · · · · · · ·
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in USSN	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ì		location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
ļ		ĺ	\			LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
						GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE
	0.760	<u> </u>	0.00			TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	Α	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR
						TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP
		ł			{	AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
		İ	ļ	<u> </u>		DL*NTSFGVIR
1213	2563	Α	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA
		l			1	TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ
		!			ļ .	IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG
		<u>L</u>	L	·	l	LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
					1	SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
						EENFGEKLHDIGFGNGFLDKT*KAQATKAKI
i					j	DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
					l	RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI
i		ľ				NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
						LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV
						EEHHLQPVQVLQTLLHSATAGTGCRRPARPP
i						PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
		1				GALGGRGGRALGGSRWPPPLPGETLFSGCKH
ļ		İ				RRRRGSDAAPGEEAGT
1218	2568	Ā	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA
						VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF
J		ł				MLDFEGEDTFHGDMAKKETVWRLE*LARLD
į			ļ !			NFEAQRALANIAADQAALEIMDMGSDYTLIP
į						NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA
.2.,	2307	1.	7002	-	204	YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYO
ļ]			LSQDPRNVWVFLATSGTLAGIMGMRFYHSG
						KL KL
1220	2570	Ā	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA
1220	2310	Λ.	1000	200	055	PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN
l						GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF
						PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI
			· '			SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL
ŀ						
1221	2571	A	9676	164	562	KERDSSTFSAAMTTMOGMEOAMPGAGPGVP
1441	ا <i>ا</i> لک	Α.	70/0	104	302	
					3	QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV
ļ				,		VQILTALMSLSMGITMMCMASNTYGSNPISV
ĺ						YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG
1000	0.600		0600	42	410	SLGMNITSS
1222	2572	Α	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF
ľ			i i			PTDENIKRKWVLAMKRLDVNAAGIWEPKKG
						DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS
·						PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	Α	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE
- 1						DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
[LKKECLCQPQKPERENLLALCCE
1224	2574	Α	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE
						GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
- 1				İ		ELVERFALYGAIEQYNALDEYPAEDFTEVYLI
						KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
						PEFETVEETRKKLQMRKAYVVKTTENKDHY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	•	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			!	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1225	2575	Α	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
						TDDVHINDISLQDYIAGVRLILL
1226	2576	Α	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT
						ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
]		i	ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
1					1	WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
						MAFVFSSLI
1227	2577	Α	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
						LAAGKMNISIDLDTNYAELVLNVGRVTLGEN
1]	NRKKMKDCQLRKQQNENVSRAVCALLNSGG
					1	GVIKAEVENKGYSYKKDGIGLDLENSFSNML
						PFVPNFLDFMQNGNYF
1228	2578	Α	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
						TLVLQKSDVEAVF
1229	2579	Α	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
1						GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
1						QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
						PLLEELGINFDHIWQKTLTVLHPLKVADGSIM
			}	j	j	NETDLAGPMVFCLAFGATLLLAGKIQFGYVY
1						GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
						GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
						WCSFSASKIFISALAMEGQQLLVAYPCALLYG
1020	0.000		1.0000			VFALISVF
1230	2580	Α	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
1						HFSPERPFMDYFDGVLMFVDISGKCKRDVCL MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
1231	2301	r.	7/44	37	1100	ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
]						LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
						LRCGWSPAEELNYTVPGPGPAGEASPROCRR
1 1			1 1			YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
						RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ
! !						SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
						NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
		·	·			WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL
i i						LVLAGVAYALPHWRWLOFTVALPNFFFLLY
		l	ŀ	,		YWCIPESPRWLISONKNAEAMRIIKHIAKKNG
						KSLPASL
1232	2582	Α	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
						YINLLNYMNWFILAGVLLDIQEVFQISDNHAG
1			'			LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS
						FGILLWSGAGLSSSFISPRYSWLF
1233	2583	Α	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
						IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
1 1						KRAYKSYVRALPLLKKMGINSILLRKSIGALE
}						VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
						FHQLV
1234	2584	Α	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
						CSRCGYGVYPAEKISCIDQIWHKACFHCEVC
1			ļ l			KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS
			•		[VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
						FHWD
1235	2585	Α	9767	52	559	IRSGAMSVDKAELCGSLLTWLQIFHVPSPCA
						SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
						SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
]						ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
	į					CARPGHSARNNTDKNLPHTAIILVTSNTYTTI
L !					/	KINFQAGRSGSCL
1236	2586	Α	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC
			ــــــــــــــــــــــــــــــــــــــ			·

1241 2591 A 9845 116 650 124 2594 A 9846 198 116 650 124 2594 A 9849 573 1620 A 9849 573 1620 A 9849 573 1620 A 9849 573 1620 A 9849 573 1620 A 9849 573 1620 A 9849 573 1620 A 9840 A 9840 198 A 9840	SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
amino acid residue of peptide residue of peptide sequence T=Threonine, V=Valine, W=Typtophan, Y=Tyrosine, X=Unknow, Y=Stop codon, /=possible nucleotide deletion, V=possible nucleotide de	seq-				correspondi		
residue of peptide pep	uence			914			
peptide sequence							Y=Tyrosine, X=Unknown, *=Stop codon,
]				-	/=possible nucleotide deletion, \=possible
1237 2587 A 9793 266 315 NILAINYEPPERJELIZOSQNPKAFALITICHE (QKINNGULPVSIDALTPELVVCLVSELTHES RYKPTREVCTIOPGGCS 1238 2588 A 9802 537 967 SELGAGRSDREAMEAAVKEEISVEDEAVDKNI RPCNKLAFYRRQKQWLSKXSTYRALLDSVT TDEDSTRFQIINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTREQUINEASKYPLLAEIYGIBONIPRILK INSETPLICIPE PROCESTRE PROC					sequence		
1237 2587 A 9793 266 515 NILAHIYFFFFLILRDSQNNKAFALT.CHI				\	1		
1238	1237	2587	A	9793	266	515	
FRDCNKIAFYRROKOWLSKKSTYRALLDSYT TDEDSTRPQINIBASK VPLLABIYGIGANIFRLK NEETPLKPRFEVPDVLTSKPSTVRLISCSGDT SSLLADGKOLKC NEETPLKPRFEVPDVLTSKPSTVRLISCSGDT SSLLADGKOLKC SSLLADGKOLKC SSPIVVLSGSMEPHRIKQLLFLINFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJLFLTNFRED PIRAGEIVYK VEGNEPHRYBKJKKMIVK VIEKDNG DIKTLYGINNEDDKGSKY POPLDPAMLLWAQGFVLEAVACQDNDDYLR YGULFEDLDCNODGVVDIBLQEGLRNWSSAF DPNSEEHG SPARKSKNRTDWITTAFKNKKMTENLAAPEA LDSSTHISSTATOSRAKMNTPAPTFSTVPANJE GGSGPPPCAPHDRYSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPFIPAVRENRN SSHSVVEFLFKRTKTPSPFIPAVRENRN TISSGPPPCAPHDRYSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTPSPFIPAVRENRN TISSGPATEPPASILSASSSDDECKEKTEDRYS LGSSLDSGMKTPLCRICFGGPEQGELLSPCRC DGSVKCTHJPQPLLIKSSKGSCUCKTKPQ LGSSLDSGMKTPLCRICFGGPEQGELLSPCRC DGSVKCTHJPQPLLIKMSERGCWSCECLCYYKY HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS LFLIASISWLIWSTFSPRARWQRQDLLFQLTVC MYGFMDVMIXAVDSEDMVQAAKEVGKRWS DJPP DTHISNIFILANQVAKGPFIVYCSDGFCELAG FARTEVMQ PROMOVEN							1
TDEDSTRQINEASKVPLLAEIYGIGGNIFRLK INEETPLKPRFEPVPVLTSRFSTVPLISCSGDT GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSLILADGKGDLKC GSESPIVVLSGSMEPAHKGDLLFLTNFRED FRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG FRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG GDKFLTKGDNNEGDDRGSYX GTGDPIPCAARRGGGGECCGAGWVAEWS POPLDPAMLLWMQGFVLEAVACQDNDVLR GUFEDLDCNGGDVVDILEQGLRNWSSAF DFNSEEHG GSGGPPCAPHDRVSSTAGDBVLR GGGGGCCGAGWVAEWS POPLDPAMLLWMQGFVLEAVACQDNDVLR GUFEDLDCNGGDVVDILEQGLRNWSSAF DFNSEEHG GSGGPPCAPHDRVSSVLQCDTQAMDHKTE SHSWVEFLIKRKTSPFIFTAVARENN LDSSTHSSSTATQSRAKMNTPAPITSTVPAIPR GGSGGPPCAPHDRVSSVLQCDTQAMDHKTE SHSWVEFLIKRKTSPFIFTAVARENN TISCGPATEPFASLLSSSASSDDFCKEKTEDRYS LGSSLDSGMRTPLCKITCQGGGLLSSPCRC LDSSTHSSTATQSRAKMNTPAPITSTVPAIPR GGSGCPPCAPHDRVSSVLQCDTQAMDHKTE SHSWVEFLIKRKTSPFIFTAVARENN TISCGPATEPFASLLSSSASSDDFCKEKTEDRYS LGSSLDSGMRTPLCKITCQGGGLLSSPCRC LGSSCCTHQPCLIKWISTERGCWSCCELCYYKY HVIAISTKNPLQWQASILTVEWQVAAALIGS LFLIASISWLIWSTFSPSARWQRQDLLFQCYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DPP GTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ THE STATE THE S	1238	2588	Α	9802	537	967	
INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT			•		ļ		, · · ·
MNKRQLYYQVLNFAMIVSSALMIWKGLINLT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVVL.GSGMEP.HRVIT GSESPIVE.CAMR.RGGGECCGAGWVAEWS PQFLDFAMRLUMQGFVLEAVACQDNDDYLR YGLFFEDL.CONGGVV.DIG.GGER.CONS.SF PQFLDFAMRLUMQGFVLEAVACQDNDDYLR YGLFFEDL.CONGGV.DIG.GGER.CONS.SF PQFLDFAMRLUMQGFVLEAVACQDNDDYLR YGLFFEDL.CONGGV.DIG.GGER.CONGWS.SF PQFLDFAMRLUMQGFVLEAVACQDNDDYLR YGLFFEDL.CONGGV.DIG.GGER.CONG.SF PARGKSNRTIDVMITAPKNKKMTENL.APEA LDSSTHSSSTATQSRAKMNTPAPIPSTVPAIPR GS8GFPP.CAPIDRVSSVLQCDQAMDHKTE SSHSVVEIL.FKRIKTFSPHAVARNRN SSHSVVEIL.FKRIKTTSPHAVARNRN SSHSVVEIL.FKRIKTTSPHAVARNRN SSHSVVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTTSPHAVARNRN FSSSVEIL.FKRIKTSPHAVARN FSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSPHATAT FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEIL.FKRIKTSPHAVARN FSSSSSSVEI							
1240 2590 A 9819 3 305 TDGRDPLPCAARRGGGGECCAGWAEWS POPLDPAMILWQGPVLEVAYACQDNDDYLR YGLIFFDLDCNODGVDIIELQEGRNWSSAF POPLDPAMILWQGPVLEVAVACQDNDDYLR YGLIFFDLDCNODGVDDIIELQEGRNWSSAF POPLDPAMILWQGPVLEVAVACQDNDDYLR YGLIFFDLDCNODGVDDIIELQEGRNWSSAF POPLDPAMILWQGPVLEVAVACQDNDDYLR YGLIFFDLDCNODGVDDIIELQEGRNWSSAF POPLDPAMILWQGPVLEVAVACQDNDDYLR YGLIFFDLDCNODGVDDIIELQEGRNWSSAF POPLDPAMILWQGPVLEVAVACQDNDDYLR YGLIFFDLDCNODGVDDIIELQEGRNWSSAF DSPARGKSNRTIDVMITAPKNKKMTENLAAPEA LDSSTINSSSTATGSRAKMNTPAPITSTYPAIPR GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SSHSVVEFLFRRTKTPSPFPAVENRN TISCGPATEPPASILSSASDDFCKERTEDRYS LGSSLDSGMRTPLCRICFQGPEGGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKNPLQWQNISLTVIEKQVQNAALIGS LFLIASISWLIWSTESPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVAAKEVGKWS DIPP DGFLYLCSAMASESSPLLAYRLIGEEGVAL PANGAGGPGGASARKLSTRLGGVVQVAAALIGS LFLIASISWLIWSTESPSARWQRQDLLPQICYG MYGFMDVMIVAVDSEDMVAAKEVGKWS DIPP DGFLYLCSAMASESSPLLAYRLIGEEGVAL PANGAGGPGGASARKLSTRLGGVVYTUSLMSTLUGVVYTUSLMST SIVVPLIGEFVVGHGGLQAAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGSIGILMFYLA NVCGCAVSLLGLVESVLDVFGA KSKCRFPEGLSEGFFWRKEALSSGVQEAE KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR VVPPLPAREVMISRTLGPEVGSIGILMFYLA NVCGCAVSLLGLVESVLDWFGA KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR VVPPLPAREVMISRTLGPEVGSIGILMFYLA NAFHSPRWGGIMYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTISLA LTVLSVCAIATNGAVGGGRYVTUSSTRSSLLPQDMM SYIGPKRTAVVRGIMHREANNIGGRAVAC LSGPTSEGLMTWELDRLLWARSVENLATATT TLTISLA PROSASSSYJLDMHSLHPVINPVESRLGSSAA SLYPVLNFLLVYPELAHSPLYIQDKGAPVAT NAFHSPRWGGIMYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELQDRLWARSVENLATATT TLTISLA PROSASSSYJLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLVPFLAGRINFYDPKSRLATATT TLTISLA PROSASSSYJLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLVPFLAGRINFYDPKSRLATATT TLTISLA PROSASSSYJLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLVPFLAGRINFYDPKSRLATATT TLTISLA PROSASSSYJLDMHSLPHVINPVESRLGSSAA SLYPVLN	1239	2589	A	9805	105	540	
PIRAGEIV PIRAGEIV			ł	Ì			
1240 2590							
1240					i		
1241 2591 A 9834 841 1209 SPARGKSNRTIDVMITAPKNKKMTENLAAPEA LDSSTHSSSTATQSRAKMNTPAPITSTTYPAIPR GSSGGPPPCAPHDRYSSVLQCDTQAMDHKTE SSHSVVEFLEKRTKTPSPFIPA VRENRN 1242 2592 A 9843 3 589 TISCGPATEPPASILSSASSDDFCKEKTEDRYS LGSSLDSGMRTPLCRICPQGPEQGELSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKNPLQWQAJSILTVIEKVQVAAAILGS LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP WRISHHAGKMPVMKGLLAPPONTFLDTIATRF DGTHSNFILANQAVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLIGEGGVAL PANGAGGGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGILQALAMILVAYFILA LTVLSVCHAIATIOAQOGAVCILQHEWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSILGILVESVLDVFGA SYGFRETAVRGIMHERALSSGSVQEAE AMIDEPQEQAEGSLTVYVISEHSSILPQDMM SYIGPKRTAVRGIMHERALSSGSVQEAE AMIDEPQEQAEGSLTVYVISEHSSILPQDMM SYIGPKRTAVRGIMHERALSSGSVQEAE AMIDEPQEQAEGSLTVYVISEHSSILPQDMM SYIGPKRTAVRGIMHERALSSGSVQEAVE KSSLGYEIITFSLLNPDKSHIDVYWDIEGAVRR YVQPFLANAIGAAGINFSVDSQILYYAMIGAND RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFILLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYMASVLPVRV EVDMVRVMEVFLAQILRILFGIAQPQLPPKCL LSGFTSEGIMTWELIALWARSVENLATATT TLTSLA LTVLSVGAGGGASCILQHEWG SARSLGRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTATT TLTSLA SESSON SARSLGRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK RREPPAGADSLSWGAGPRISSYY FVRNKKMTRSCSAVGGSTRDTVLSRERGLSF HQPFTDTIQRSK WIRAVNRV DPTSRKKIWIPGP GAILCSKHQESDFESYGIRKKLKKGAVPSVS SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK RREPPAGADSLSWGAGPRISSYY SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK RREPPAGADSLSWGAGPRISSYY SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK RREPPAGADSLSWGAGPRISSYY SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK RREPPAGADSLSWGAGPRISSYY SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALKAKAVNRVDFTA SARSLGRRPRIAMVTVGNYCEAEGVGPAWM QDGLSPCFFFTLVPSTRMALGTLAVILAUPCK GABLCKHAR	1240	2590	Α	9819	3	305	L '
DPNSEEHG							
1241		1					
LDSSTHSSSTATQSRAKMNTPAPIPSTVPAIPR GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE	1241	2501		0834	941	1200	
1242 2592 A 9843 3 589	1241	2391	^	7634	041	1209	
1242							
LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKNPPLQWQAISLTVIEKVQVAAAILGS LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP		<u> </u>					
DGSVKCTHQPCLIKWISERGCWSCELCYYKY	1242	2592	A	9843	3	589	
HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS			1				
LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP 1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVYPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQLYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLVIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRIPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGISF HQFFTDTLQRSK WIRAVNVDPRSKKIWIPOPG GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		!	1	ļ			
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGILQALAMILVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWRAVNVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			1				LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VPVLPAREVMISRTLGPEVGGSIGIAMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK REEPPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			İ				,
1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMYYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGFAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HOFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1242	2502	_	0846	108	411	
PARTEVMQ	1243	2393	^	7040	170	411	
PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRPPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFFTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGFMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1244	2594	Α	9848	116	650	t e
LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		ŀ	ł				f .
VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV RRERPAGADSLSWGAGPRISSYV FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTTQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		ļ					
1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			1				
AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS					ļ		
SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	2595	A	9849	573	1620	
AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			[1			
RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				1			
NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS]]	}			
EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS]			l i
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			ļ] ,		}	1
ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1246	2506		0850	114	464	
QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1240	2370	^	7670	114	704	` `
1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				1			
HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1247	2597	Α	9851	2	327	1 ' '
			1		1		LYKVFKYSSRCTS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RVDDFVYSKGGKDAGGADVSLACRRQSIPEE
1240	2390		7033	30		FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	Α	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARIHAPRRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	A	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV QLA
1251	2601	A	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	Α	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL
1255	2605	Α	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT
1256	2606	A	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	Α	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence.	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH QRRGPSCGASGDPQCVGSPHPQRARPLLARP GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS YLGECGSSSYVTGAACISPVLRCREWFEAGLP WPYERGFLLHQKIALSRYATALEDTVDTSRL FRSRSLREFEEALFCHTKSFPISWDAYWDRND PLRDVDEAAVPVLCICSADDPVCGPPDHITLTT ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR ERPAARSGPEMRVRYPVVAAVLAPYLALSQD PMYKSSASGQGASGSYNHVREEMLIKAGGA MSRRVVRQSKFRHVFGQAAKADQAYEDIRV SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL PLAK
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL PPRPGRSHRKRKLVSTK
1260	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE KKQKASQNLVVLAREDAGAEKIFRSNGVQLL QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRTEGGPGAGSG RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF PTRVDHNGALLAFSPPPPQRQRRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVRQFNIYDY
1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	Α	10005	2	209 ·	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD LQLRNLSVADHSKTQVQKKENKSLKRDTKAI IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS QPKLDRTSSFRQILPRFRSADHDRYRGWSMW DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV SAPRRAASGPSGSAPAVAAAAAQPGSYPALS AQAAREPAAFWGPLARDTLVWDTPYHTVW DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS PESVALIWERDEPGTEVRITYRELLETTCRLA NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA CARIGAVHTVIFAGFSAESLAGRINDAKCKVV ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH VLVAHRTDNKVHMGDLDVPLEQEMAKEDP VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT QAGYLLYAALTHKLVFDHQPGDIFGCVADIG WITGHSYVVYGPLCNGATSVLFESTPVYPNA GRYWETVERLKINQFYGAPTAVRLLLKYGD AWVKKYDRSSLRTLGSVGEPINCEAWEWLH RVVGDSRCTLVDTWWQT
1274	2624	A	10017		3750	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG KTLGSFFGSLPGFSSARNLVANAHSSARARPA ADPTGAPAAEAAQPQAQVAHPEQTAPWTE KELQPSEKMVSGAKDLVCSKMSRAKDAVSS GVASVVDVAKGVVQGGLDTTRSALTGTKEV VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTTFGSGVTSAVNVAKGAA QTGVDTAKTVLTGTKDTVCSGVTGAAN

NO. of peptide sequence NO. of peptide sequence NO. of peptide sequence No. of peptide No. of pep	SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucle- cotide seq- uence Seq- uence				`			
cottide sequence uence uence				1			
sequence Sequence Uence 99446 914 914 916 915 914 914 914 914 914 914 914 914 914 914 914 916 915 914 915			ł				
uence 914 ng to first amino acid residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Tytytophan, Y-Tyrosine, X-Uaknown, *-Stop codon, Y-possible nucleotide deletion, mostible nucleotide deletion, \mathbb{mostible nucleotide nu			}				
amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence sequ			1				
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, >-possible nucleotide insertion vakGAlQGGLDTTKSVLTGTKDAVSTGL. VAKGARQGGLDTTKSVLTGTKDAVSTGL. VAKGARQGGLDTTKSVLTGTKDAVSTGLTGANVAKGAVQTGLDTTKSVLTGTKDAVSTGLTGANVAKGAVQTGLDTTKSVLTGTKDAVSTGLTGANVAKGAVQTGLDTTKSVLTGTKDAVSTGLTGANVAKGAVQTGLDTTKSVLTGTKDAVSTGLTGANVAKGAVQTGLDTTKSVLTGTKDAVSAGLMGSGTVLTGTAVAKGAVQTGUDTTKSVLTGTKDAVSAGLMGSGTVLTGTAVAKGAVQTGUDTTKSVLTGTKDAVSAGLMGSGTVLTGTTGANVAKGTVQTGVAKAKGAVQTGLDTKTVLTGTKDAVSAGLMGSGTTGATHTGLSTFONULPSTPATSWGGLTSS TDNGGGGTALSPQEAPFSGISTPPDVLSVGAWAAATTKGLATDVATTFOGAAPGREE LLATTHGPEEAPRLAMLONELEGLGDIFH NAEEQAQLAASQPGEPKVLSAEQGSVFVRI LQPSFRQAFFHAVSHLQHGQFQARDTL/QDCFRL VAKEAAATTKVEKKKEVLAPVTKPV VDKNGGTRVVKLPTMPRYPTEDVPRKLLKKFPS VKRKYSAANTKVEKKKKKEVLAPVTKPV VKRKYSAANTKVEKKKKKEVLAPVTKPV VKRKYSAANTKVEKKKKKEVLAPVTKPV VKRKYSAANTKVEKKKKKEVLAPVTKPV VKRKYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSAANTKVEKKKKKEVLAPVTKPV VKRYSTAGTGSTGLTSCAPCVRRLFF CRIPNSLPYFHKRPQARMLLLALFCVAV VWGVFRIEDQ VKRYSEQFTCPNFRGGGGLFS VSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDQIKRENINGSITMSGHVCFI PPQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VGLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VGLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VGLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VGLKLTINTVGFGFONLGVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VGLKLTINTVGFGFONLGVGEGGG VGLCCHEVEKYKTFNKYWCC CLPIWHEMVFGGSEOVKSDQVITTDHR TTTVTLENLTADDAGKYRCGIATILOEDGI		į.			, –		
peptide /-possible nucleotide deletion, \-possible nucleotide insertion vakGaiqGGIDTTKSVLTGTKDAVSTGLT VakGaiqGGIDTTKSVLTGTKDAVSTGLT VakGaiqGGIDTTKSVLTGTKDAVSTGLTGKAAVAKGAVQMGVDTAKTVLTGTKDAVCS TGAANVAKGAVQMGVDTAKTVLTGTKDAVCS GTKDAVSTGALTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAMSMAK VQGLDTTKTVLTGTKDAVSAGLMGSGG TGATHTGLSTFQVTGVTGAVNLAKGTVQTGAMSMAK VQGLDTTKTVLTGTKDAVSAGLMGSGG AWEAAATTKGLATDVATTGQAAPGREE LLATTHGPEEAPRLAMLQNELEGI.GDIFH NAEEQAQLAASQPGFKVLSAEQGSYPVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLAQDCFRL VLTGTKTGATATTAVATAVATAVATAVATAVATAVATAVATAVATA		1	[1		
		l	1	1		5040000	
VAKGAIQGGLDTTKSVLTGTKDAVSTGL VKLAKGTVQTGMDTTKTVLTGTKDAVCS TGAANVAKGAVQMOUDTAKTVLTGTKDAVCS CSGYTGAANVAKGAVQGGLBKTTQNIATG NTLGSGYTGAANVAKGAVQGGLDTTKSS GTKDAVSTGLTGAVNLAKGTVQTGVDTS VL1GTKDTVCSGVTGAVNVAKGTVQTGVDTS VL1GTKDTVCSGVTGAVNVAKGTVQTGVDTS VL1GTKDTVCSGVTGAVNVAKGTVQTGVDTS VL1GTKDTVCSGVTGAVNVAKGTVQT AKTVLSGAKDAVTLSGTGVANVAKGTVQT VDASKAVLMGTKDTVFSGVTGAMSMAK VQGGLDTTKTVLTGTKDAVSAGLMGSGS TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPPSGISTPPDVLSVG AWEAAATTKGLAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK YKRKYSAANTKVEKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL KKPPS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSI ASRPQSPTTPWCLPRFYMKHKRDDGFGE EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPP CRIPNISLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVFLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEJKRENIRSLTMSGHVGF PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNEEDVESSHFCPNVKLKAQTTELQC VQLKLTIVNTTVGFGDQNIKEENGLGSMSS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQPYLIPNSPPHRLADVADVHLE LSGAFGCYHLDVTNNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCPPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWC CLPIWHEMVETGGSGEVVXGQVHTDHPC TTTVTLENLTADDAGKYRCGIATILQEDGI			1		1		
VKLAKGTVQTGMDTTKTVLTGTKDAVCS TGAANVAKGAVQMGVDTAKTVLTGTKDAVCS TGAANVAKGAVQMGVDTAKTVLTGTKD CSGVTGAANVAKGAVQTGLDTTKSV GTKDAVSTGLTGAVNIAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS AKTVLSGAKDAVTTGVTGAMSMAK VQGGLDTTKTVLTGTKDAVSAGLMGSGN TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPPSGISTPPDVLSVQ AWEAAATTKGLATDVATFTQGAAPGREE LLAITHGPEEAPRLANLQNELEGLGDIFH NAEEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRXYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPRVPSCSRVFCPVPPGGCGLPSI ASRRQSPTTPW-LPRRYMKHKRDDGPEK, EAVDVTPWTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGVSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLAFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVEKLAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SGGSAQOPYLPINSPPHRLADVADVHLF, LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKSLRPCCPRDFHAGCCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCI CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI		 -		 	boquesto		
TGAANVAKĞAVQMGVDTAKTVLTĞTKD CSGVTGAANVAKGAVQTGLKTTQNIATG NTLGSGVTGAAKVAKGAVQTGLKTTQNIATG NTLGSGVTGAAKVAKGAVQTGLKTTQNIATG NTLGSGVTGAAKVAKGAVQTGLDTTKSV GTKDAVSTGLTGAVNLAKGTVQTGVTS VLTGTKDTVCSGVTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGVTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGTTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGTTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGTTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGTTGAVNVAKGTVQTGV AKTVLSGAKDAVTTGTGTAVNAKGTVQTGV AWEAAATTKGLATDVATFTQGAAPGREI LLATTHGPEZAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGPKVLSAEQGSVFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLLL KKPPS 4 SGSLRFSPPRVPSCSRVFCPVPPGGCGLFS ASRPQSPTTPWCLPRYMKHKRDDGPEK EAVDVTPVMTCVFVVMCCSMLVLLYFY LVYVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCPNILCVGETGIGKSTL LFNTNFEDYESSHCPNVLKLAQTYELQE VQLKLTIVNTVGFGDQNKEERQLGRSQS PQKYRSEQIHVEPKKCTSFWKGALGKWA SSQQSAQQPYLPINSPPHRLADVADVHLIF- LSGAFGCYHLDVTVNFEKKQQNRDEQEG VGLKLTIVNTVGFGDQNKEERQLGRSQS PQKYRSEQIHVEPKKCTSFWKGALGKWA SSQQSAQQPYLPINSPPHRLADVADVHLIF- LSGAFGCYHLDVTVNFEKKQQNRDEQEG VGLKGAGGSUSVGAGPSUSPACTTOSGP VMGAVGESLSVQCRYEEKYKTTNKYWCI CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI		j	}]]	
CSGVTGAANVAKGAVQTGLKTTQNIATG NTLGSGVTGAAKVAKGAVQGGLDTTKSV GTKDAVSTGLTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS AKTVLSGAKDAVTTGVTGAVNVAKGTVQ VDASKAVLMGTKDTVFSGVTGAVNNAKGTVQ VDASKAVLMGTKDTVFSGVTGAMSMAK VQGGLDTTKTVLTGTTKDAVSAGLMGSGN TGATHTGLSTFQNWLPSTPATSVGGLTSS TDNGGEQTALSPQEAPFSGISTPPDVLSVG AWEAAATTKGLATDVATFTQGAAPGREI LLATTHGPEEAPRLAMLQNELEGLGDIPH NAEEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHQQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSKK YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL KKPFS ASSRQSVTTFWCLPRTYMKHKRDDGPEK ASSRQSVTLFWCLPRTYMKHKRDDGPEK EAVDVTPVMTCVFVVMCCSMLVLLYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGYFRNEDQ TCMSSQGDDEQIKRENIRSLTMSGHVGF PDQLVNRSIQQGFCPNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIHVVEPKKCTSFWKGALCKWA SSQGSAQQPYLPINSPPHRLADVADVHLIF LSGAFGCYHLDVTVNEFKKQQNRDEQEG VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPNVELKAQTYELQE VGRYBSSHFCPRDFHLAGCTVVSGP VMGAVGESLSVQCRYEEKYKTNKVWCI CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADDAGKYRCGIATILQEDGI VFTYTLENLTADD					1	1	
NTLGSGYTGAAKVAKĞAVQGGLÖTTKS\ GTKDAVSTGLTGAVNLAKGTVQTGVDTS\ VLTGTKDTVCSGVTGAVNVAKGTVQ\ VLTGTKDTVCSGVTGAVNVAKGTVQ\ VDASKAVLMGTKDTVFSGVTGAVNVAKGTVQ\ VDASKAVLMGTKDTVFSGVTGAVNVAKGTVQ\ VDASKAVLMGTKDTVFSGVTGAVNVAKGTVQ\ VDASKAVLMGTKDTVFSGVTGAVNVAKGTVQ\ VDASKAVLMGTKDTVFSGVTGAVNSAGLMGSG\ TGATHTGLSTFQNWLPSTPATSWGGLTS\ TDNGGEQTALSPQEAPFSGISTPPDVLSVG\ AWEAAATTKGLATDVATFTQGAAPGREI\ LLATTHGPEAPRLAMLQNELEGLGDIFH\ NAEEQAQLAASQPGKVLSAEQGSYFVRI\ LGPSFRQRAFEHAVSHLQHGQFQARDTLA\ QDCFRL \ LGPSRRQRAFEHAVSHLQHGQFQARDTLA\ QDCFRL \ TILARKKEKTCPCKKEIGRNSRSGMYSRK\ YKRKYSAANTKVEKKKKEVLAPVTKPV\ DKNGGTRVVKLPTMPRYYPTEDVPRKLL.\ KKPFS \ SGSLRFSPFRVPSCSRVFCPVPPGGCGLFS\ ASRQSPTTPWCLPRYMKHKRDDGPEK\ EAVDVTPVMTCVFVVMCCSMLVLLYYF\ LVYVVIGICLASATGLYSCLAPCVRRLPF\ CRIPNISLPYFHKRPQARMLLLALFCVAV\ VWGVFRNEDQ \ LFVVNVIGICLASATGLYSCLAPCVRRLPF\ CRIPNISLPYFHKRPQARMLLLALFCVAV\ VWGVFRNEDQ \ LFVVNSIQQGFCFNILCVGETGIGKSTL.\ LFNTNFEDVESSHFCPNVKLKAQTYELQE\ VQLKLTTVNTVGFGQDINKEERQLGRSQS\ PQKYRSEQHPVEPKKCTSFWKGALGKWA\ SSGQSAQQPYLPINSPPHRLADVADVHLF\ LSGAFGCYHLDVTVNEFKKQQNRDEQEG\ KGDQEQGSWKHGADPLRGGEM \ SGGSAQGSVYGEKYKTFNKYWCI\ CLPIWHEMVETGGSEGVVRSDQVIITOHPC\ CLPIWHEMVETGGSEGVVRSDQVIITOHPC\ CLPIWHEMVETGGSEGVVRSDQVIITOHPC\ CLPIWHEMVETGGSEGVRSDQVIITOHPC\ CLPIWHEMVETGGSEGVRSDQVIITOHPC\ CLPIWLEMVETGGSEGVRSDQV		ł					
GTKDAVSTGLTGAVNLAKGTVQTGVDTS VLTGTKDTVCSGVTGAVNVAKGTVQTGVDTS AKTVLSGAKDAVTTGVTGAVNVAKGTVQTGAVNVAKGTVQTGAVNVAKGTVQTGAVVAKGTVQTGAVNVAKGTVQTGAVPAKGTVQDAVAKGTVQDAVAKGTVQDAVAKGTVQDAVAKGTVQDASKAVLMGTKDTVFSGVTGAMSMAKQDGGATSTATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPFSGISTPPDVLSVGAWEAAATTKGLATDVATFTQGAAPGREILLATHTGPFEAPTLAMLQNBLEGGLGDIFH NAEEQAQLAASQPGPKVLSAEQGSYFVRILGPSSRQAFEHAVSHLQHGQFQARDTLAQDCFRL 1275		ł	1		ł	}	
VI.TGTKDTVCSGVTGAVNVAKGTVQTGV	1)	1	1	!]	
AKTVLSGAKDAVTTGVTGAVNVAKGTVV VDASKAVLMGTKDTVFSGVTGAMSMAK VQGGLDTTKTVLTGTKDAVSAGLMGSGN TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPPSGISTPDVLSVG AWEAAATTKGLATDVAJTTQGAAPGREI LLAITHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGFKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL LGPSFRQRAFEHAVSHLQHGQFQARDTLA VKRYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLLI KKPFS LSGAFTVVLLPTMPRYYPTEDVPRKLLI KKPFS LYVVIGIFCLASATGLYSCLAPCVRRLIPF CRIPNISLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ LYVVIGIFCLASATGLYSCLAPCVRRLIPF CRIPNISLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ VSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGF PPQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYPELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGGSWKHGADPLRGGEM KGDQEGGGSWKHGADPLRGGEM LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEGGGSWKHGADPLRGGEM KGDQEGGGSWKHGADPLRGGEM VMGAVGESLSVQCRYEEYYRTFNKYWCC CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI			ļ			}	
VDASKAVLMGTKDTVFSGVTGAMSMAK VQGGLDTTKTVLTGTKDAVSAGLMSGMS TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPFSGISTPPDVLSVG AWEAAATTKGLATDVATFTQGAAPGREE LLATTHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGFKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK YKRKYSAANTKVEKKKKEVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLFS! ASRPQSPTTPWCLPRRYMKHKRDDGPEKE EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGYFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYPELQE VQLKLTIVTVGFGDQINKEERQLGRSQS PQKYRSEQHPVEPKKCTSFWGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQG KGDQEQGGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCG CLPIWHEMVETGGSEGVVRSDQVIITDHP TFTVTLENLTADDAGKYRCGIATILQEDGI		1					1
VQGGLDTTKTVLTGTKDAVSAGLMGSGN TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPFSGISTPDVLSVG AWEAAATTKGLATDVATFTQGAAPGREI LLAITHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLLI KKPPS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPS ASRPQSPTTPWCLPRRYMKHKRDDGPEK EAVDVTPVMTCVFVVMCCSMLVLLYYF LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQSDDEQIKRENIRSLTMSGHVGF PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDVESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSQQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQNRDQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVYRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI		ì	l		ł		,
TGATHTGLSTFQNWLPSTPATSWGGLTSS TDNGGEQTALSPQEAPFSGISTPPDVLSVG AWEAAATIKGLATDVATFTQGAAPGREE LLAITHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVKLPTMPRYYPTEDVPRKLL: KKPPS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPS ASRPQSPTTPWCLPRRYMKHKRDDGPEK EAVDVTPVMTCVPVVMCCSMLVLLYYFY CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRTTYPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGQQINKEERQLGRSQS PQKYRSEQHPVEPKKCTSFWKGALGKWA SSGQSAQQPYLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCI CLPIWHEMVETGGSEGVVRSDQVIITDHPI				l	i		
TDNGGEQTALSPQEAPFSGISTPPDVLSVG AWEAAATIKGLATDVATFTQGAAPGREE LLA'ITHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASQPGRVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPS ASRPQSPTTPWCLPRRYMKHRDDGPEK EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHPVEFKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLPS LGGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWC CLPIWHEMVETGGSEGVVRSDQVIITDHPY TFTVTLENLTADDAGKYRCGIATILQEDGI		İ		}	ļ		
AWEAAATTKGLATDVATFTQGAAPGREI LLA'ITHGPEEAPRLAMLQNELEGLGDIFH NAEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEK EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSiQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPHRLADVADVHLPS LGGAFGCYHLDVTVNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHLAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWC CLPIWHEMVETGGSEGVVRSDQVIITDHP TFTVTLENLTADDAGKYRCGIATILQEDGI		}			1		1
LLATTHGPEEAPRLAMLQNELEGLGDIFH NAEEQAQLAASOPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEKG EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGGI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCE CLPIWHEMVETGGSEGVVRSDQVIITDHPM TFTVTLENLTADDAGKYRCGIATILQEDGI		1		1			
NAEEQAQLAASQPGPKVLSAEQGSYFVRI LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL			ł	l	ſ		
LGPSFRQRAFEHAVSHLQHGQFQARDTLA QDCFRL 1275 2625 A 10025 124 415 TILARKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGGLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEKG EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLLFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPM TFTVTLENLTADDAGKYRCGIATILQEDGI			l	ł .	ł		
DDCFRL 1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL. KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEKGEAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLLFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPGTTTVTLENLTADDAGKYRCGIATILQEDGI		1	i	Ì	ł		
1275 2625 A 10025 124 415 TILARKKEKTCPCKKEIGRNSRSGMYSRK. YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPS! ASRPQSPTTPWCLPRRYMKHKRDDGPEK! EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLLFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS. PQKYRSEQHPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCP VMGAVGESLSVQCRYEEKYKTFNKYWCP CLPIWHEMVETGGSEGVVRSDQVIITDHPR TFTVTLENLTADDAGKYRCGIATILQEDGI		1		ļ			
YKRKYSAANTKVEKKKKEKVLAPVTKPV DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS	1000	2505		10005	100		
DKNGGTRVVKLPTMPRYYPTEDVPRKLL: KKPFS 1276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEKI EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQQYPLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	12/3	2625	A	10025	124	415	
I276 2626 A 10030 3 507 GGSLRFSPPRVPSCSRVFCPVPPGGCGLPS ASRPQSPTTPWCLPRRYMKHKRDDGPEKG EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGF PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPNSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM KGDQEQGSWKHGADPLRGGEM RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI			ľ		.	•	
1276 2626 A 10030 3 507 GGSLRFSPRVPSCSRVFCPVPPGGCGLPSI ASRPQSPTTPWCLPRRYMKHKRDDGPEKG EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQSI PQKYRSEQIHPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI			ļ	ļ	l		
ASRPQSPTTPWCLPRRYMKHKRDDGPEKG EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGGTGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	1056		<u> </u>		<u> </u>		
EAVDVTPVMTCVFVVMCCSMLVLLYYFY LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	1276	2626	[A	10030	3	507	
LVYVVIGIFCLASATGLYSCLAPCVRRLPF CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQSY PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI					ļ		
CRIPNNSLPYFHKRPQARMLLLALFCVAV VWGVFRNEDQ 1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI		1	l	ļ			
1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI				1			
1277 2627 A 10035 51 869 YSRFTVPLPATMASSEVARHLLFQSHMAT TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPG TFTVTLENLTADDAGKYRCGIATILQEDGI							
TCMSSQGSDDEQIKRENIRSLTMSGHVGFI PDQLVNRSIQQGFCFNILCVGETGIGKSTLI LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQIIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							
PDQLVNRSIQQGFCFNILCVGETGIGKSTLL LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	1277	2627	A	10035	51	869	
LFNTNFEDYESSHFCPNVKLKAQTYELQE VQLKLTIVNTVGFGDQINKEERQLGRSQS: PQKYRSEQIEPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							
VQLKLTIVNTVGFGDQINKEERQLGRSQS PQKYRSEQHIPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							
PQKYRSEQHPVEPKKCTSFWKGALGKWA SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							
SSGQSAQQPYLPINSPPHRLADVADVHLFS LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI		[VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN
LSGAFGCYHLDVTVNEFKKQQNRDEQEG KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	ı	i i					PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE
KGDQEQGSWKHGADPLRGGEM 1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							SSGQSAQQPYLPINSPPHRLADVADVHLFSSV
1278 2628 A 10036 3 457 RAFDVRRKKSLRPCCPRDFHAGCLTVSGP VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI		<u> </u>]			LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS
VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI							
VMGAVGESLSVQCRYEEKYKTFNKYWCF CLPIWHEMVETGGSEGVVRSDQVIITDHPC TFTVTLENLTADDAGKYRCGIATILQEDGI	1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST
TFTVTLENLTADDAGKYRCGIATILQEDGI							VMGAVGESLSVQCRYEEKYKTFNKYWCRQP
TFTVTLENLTADDAGKYRCGIATILQEDGI	.						CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL
							TFTVTLENLTADDAGKYRCGIATILQEDGLSG
	_						FLPDPFFQVQVLVSSASSTENSVKTP
	1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR
		!					SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA
QMPCMWWYPFWG							
	1280	2630	Α	10043	2	344	RATWHNAGKEREAVOLMAGAEKRVKASHS
							FLRGLFGGNTRIEEACEMYTRAANMFKMAK
					i		NWSAAGNAFCQAAKLHMQLQSKHDSATSFV
DAGNAYKKADPQGKTARHVACYLCV		j					l
	1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG
		2051	**		J., U	0.0	PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
	1282	2632	_	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY
	. 202	2002	^	10004		1040	
]		NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA
							ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF
							YEFQLTAVSEGGVLSESSSTANITVVASDSPY
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1							GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR
1 1 1 1 1	ļ						LWYKTMSGTAEAGLDFVPAAGELLFEAGEM
							RKSLHVEILDDDYPEGPEEFSLTITKVELQGR
							GYDFTIQENGLQIDQPPEIGNISIVRIIMKNDN
AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLF		L					AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Głutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				Sequence		YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA
						NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII
						LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	Α	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV
						TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA
						KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE
1285	2635	A	10092	290	728	LP KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP
						AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL
						IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	Α	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL
						ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR
						VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	Α	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD
						SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG
1289	2639	A	10113	237	438	KT LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT
					₩ ⁻	DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	A	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN
				į		PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI
						KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV
						LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG
						SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM
						LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP
		<u> </u>				MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				<u>}</u> .		LGNVLTSTPNAKTVNGKAESSDSGAESEEEE AC
1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL CGRGIRDSARMCSTCACVEYYGKALEVLVK GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV CTDVMARGIDIPEVNWVLQYDPPSNASAFVH RCGRTARIGHGGSALVFLLPMEESYINFLAIN QKCPLQEMKPQRNTADLLPKLKSMALADRA VFEKGMKAFVSYVQAYAKHECNLIFRLKDL DFASLARGFALLRMPKMPELRGKQFPDFVPV DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN EGRRKFIKNKAWSKQKAKKK
1294	2644	,	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF GVLFALCFSCLLAHASNLVKLVRGCVSFSWT TILCIAIGCSLLQIILATEYVTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFMALTFFVSKAT FCGPCENWKQHGRLIFITVLFSIIIWVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA
1295	2645	Α	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV
1296	2646	Α	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	А	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT
1299	2649	A	10161	1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV
1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV
1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

NO. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of muci- soid sequence No. of peptide	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence unce unce to dide tequence unce to the contraction of the contra							
South Sout			1100				
uence ue						l .	
uenice 1914 mg no first minion acid residue of peptide pe	1						
amino acid residue of peptide requence peptide sequence peptide sequence peptide sequence peptide sequence per sequence pe		dence			•	I .	
residue of peptide sequence Poptide in culceolité delicion, "possible nucleotide (inclion, "	uence			914			
peptide sequence Possible nucleotide deletion, = possible	l]					
sequence nucleotide insertion	1	Į.				sequence	
KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS KEKPERKTSPEDKJSVKHKYKGDOKOMHKTO DETELISSEKGLK VEENIQKOSQOTKLSSDDK TERSKHRPERKJSVLKGROKPVSVEHIKTDE NVRKENNKKERRLSAEKTAEHKSRRSSDSS (QKDSLGSKOHGTILORGENSSVSEDKOMDST NMDSNLKPESVVHKERKRTKSLLEEKLVLKS SKYTOKGVVKVVETELDGGATKQATTEKPD KERNTEENDSIKQRKSKVEDKPFEETGVEPV LETASSSAHSTOKDSSHEKARKNSSLMEKLLS DEKSATSTENDSKORKSKVEDKPFEETGVEPV LETASSSAHSTOKDSSHEKARKNSSLMEKLLS ENKSDDKDGKEVDSSHEKARKNSSLMEKLLS ENKSDDKDGKEVDSSHEKARKNSSLMEKLLS SRLCENRROSLSQEMAKGERLAANTLSTP SGSSLQRRKKSGDMTLIPRQFPMEIDSEPOVE NVFVSKTQDKRNNSHSQDDISSIMKOKTS ATVOKDELRTCTADSKATAPAYKFORGTOV NSMSEKHADHRSTLTKKMINGSAV SKWINDER KEPHRGTTEVNIDSETVERMLLSAPSENDEV QKNLKNTAAEEDVAGODATLEISTNILDSSPS LSSVTVVELRESVDFDVIPLFDKKTVLEGSTA STSPADHSALPNOSLTVERSEVLKTSDSKEGG GKVIMPLGKSLTGVVENEMIKEGGLVOMA KKENDLNAEPHLKQTIKKATVENGKKOGIAVD HVGLINTENYAETVKHENSTRCKVIDSID VERRINENSEVDTSAGSGSARSVLHORNOGITE DVATGFRAETSVSHALPTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT HSRNNILSTERGSEVALPCTSIEADEGLIGT GFAESETELTSTKEGESGECAVESEDRAADL LAVIAVKIEANVINSVVTEEKDAATSAGSEE KCDGSLSRDSIGVEGTTITISEVESDGAVTDEG GFAESETELTSTKEGESGECAVESEDRAADL LAVIAVKIEANVINSVVTEEKDAATSAGSEE KCDGSLSRDSIGVEGTTITISEVESDGAVTDEG GGSAVTSTGTTEGGEGFANSCTGSEDSSEGFALIS SESEENGESAMDSTVAKEGTNIPLAAGPCD DEGIVTSTGAKEEDEEGDVYTSTGRONEIGH ASTCTILGESGESCALVESSENGSGORVINDTE GGSAVTSTGTTEGGEGFANSCTGSEDSSEGFALIS SESEENGESAMDSTVAKSGTNIPSFEKEDEDITISTE PURGGFRYTTEFFARMFANNOSMSGTEKGSKDT DICSSKGRUESSTVASGGOGSTYNOFT GGSAVTSTGTTEGGEGFANSCTGSEDSSEGFALIS SESEENGESAMDSTVAKSGROVETYFOGGE GPMTSAASDQSBSQLEKVEDTTISTGLVGGS YDVLVSGEVPECVAHTESPEKEDEDITISTE NMEGTEVTTEFFARMFANNOSMSGTEKGSKDT DICSSKGROESSTANDSTANGEREPLENDATADHROEN GGKEPFOFVANSTEEGHANSPSKEREPLAADL LEARANGEREPLENDATADHROEN GGKEPFOFVANSTEEGHANSPSKEREPLAADL HEABCAGAMMANENNYDSMSGTEKGSKDT DICSSKGROESSTANDSTANGEREPLENDATADHROEN GGKEPFOFVANSTEEGHANDSTANGERGANDGT PAAVCAEEGERENVILISTAGGESEGANDSTANE	1	·			peptide		
KEKPEREKTPSEDKLSVKHHYKGDOMHIKTO DETELISSEKGIA VEREINIGKOSQOTKLSSDIK TERKSKHRNERKI SVI.GKDGKPVSEVIIKTDE NYRKENNIKKERILS SEKTKAPHIKRSRSDSK JOKDSLGSKOHGTILORRSESYSEDKCDMDST NADSSLAFERVENTEKERSTSDIK LOMDST NADSSLAFERVENTEKERTISLLEGLAVIK KSKTOCKOVKVVETELQEGATKOATTPKPD KERNTELERLADGHKSRSLKRKTSLLEKLEVLJKK KSKTOCKOVKVVETELQEGATKOATTPKPD KERNTELERLADGHKSRSLKHSKODICKEN DIKTORY LETASSSAHSTORDSSRFARKLPLAKEKYKSD KOSTSTRLERKLSDGHKSRSLKHSSKODICKEN ENKSDEMORE VERSEKARONSSLMERKL SRRLCENRROSLSQEMARGERIK LANTLSTT SOSSLORFRKSGDMTHIPEOGFMEIDSEFOYE NYFEVSKTODNRNINNSHODIDSENMKOKTS ATVOKDELRICTADSKATAPAYKRGGTOV NSINSEKHADHRSTLTKKMHIGSAVSKMPGGE KEPHRGTEVNINDESTVIRMLLASRSHODRY OKNILKNTAAEHVAQGDATLEHSTNLDSSFS LSSVTVYPLRESYPDPVIPTPKRYLLEGSTA STSPADHSALPNOSLTVERSEVLKTSDSKEGG GEGFTVYTTPTAKASTISKRHPEAHQATLLOGKQ GKVIMPLOSKLTGTVENESHLKTSDSKEGG GEGFTVTTTTAKASTIKKRHPEAHQATLLOGKQ GKVIMPLOSKLTGTVENESHLKTSDSKEGG GEGFTVTTTTAKASTIKKRHPEAHQATLLOGKQ GKVIMPLOSKLTGTVENESHLKTSDSKEGG UPDATOPTRAKASTIKRATIPEAHQATLLOGKQ GKVIMPLOSKLTGTVENESHLKTSDSKEGG UPDATOPTRAKSTIKATVENKOKOGIAVD HVVGLNTEKYABTVKLKHRRSFGKVKDISID VERNINENSEVTISTAGSGASPSVLHORNOGTE DVATOPRRAEKTSVATSTEGKOKOTVLSPVK AGAPATTTSSETROSEVALPCTSIEADGGLIGT HSRNNPLHYGARASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGGSGEGAVASEGGIAVTSA TERRAGISSSEEVDGSOGGMMMRMGYKKETEG UTCTGAEGRSDNPVTTISTGCHSOMTOTS GFAESETTLTSTKCGSGEGAVATSEGRAVADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KOCGLSSRDSGIVGTTTEE GFAESETTLTSTKCGSGEGAVTSAGGERMADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KOCGLSSRDSGIVGTTTISTGEVESDGAVTSAGG TUTCTGAEGRSDNPVTTISTGCHSOGGIGTVVEH VERAGABGMANDAPHAGGSTTISTGCROEGH ASTCTGLOESEGGVLCSVTGAOPREERMYT GAGVULGDNDAPPGTSASQEGGSVNDGTE GESAVTSTGTAGEGGDAVTSAGGTER UTCTGAEGRSDNPVTTISTGCLOCG GPWTSAASDQGBKWENTUPLVAAGPCD DEGVTSTTGTGTEGGEGDAVTSTGGEGHAASSTEAGER NEGEGEGAVTGGEGGDAVTSAGGGEGGAVTSAGG TUTCTGAEGRSDNPVTPLAGGGGTAVTSAGGGGGGGAANSFANDORSHGTERAGGGGGAAGGGGGGGAANTSAGGGGGGGAANGGGGGGGAAGGGGGGGGGG		l		L	sequence		nucleotide insertion
DETELHSSEKGLKVEENIOKOSOOOTKLSSDDK TERKSKINENERSLSVLGKOGROPSVIIKTDE NVRKENNIKERRLSAEKTKAEHKSRESSSS NVRKENNIKERRLSAEKTKAEHKSRESSSS NVRKENNIKERRLSAEKTKAEHKSRESSSSS NOKDOSLOSKOHORITORRSESYSEDKCOMDIST NMDSNILKPEVVIKERRRTKSLLEEKLVLKS KSKTOGKOVXVVETELOGGATKATTEKPD KEKNITEENDSEKORKSKVEDKPPETGOVEPV LETASSSAHSTOKDSSHRAKUPLAKEKYKSD KOSTSTELERKLSDGHKSRSLHSKDIKKKL DIKSSDMLOGKEVDSSHEKARDSSLMEKKL DIKSSDMLOGKEVDSSHEKARDSSLMEKKL SRRLCENRRGSLSOMAKGEKLAANTLTST SGSSLORPKKSGDMITLIPEQEPMEIDSEPOVE NVFEVSKTODINRNINSHODDSEPOKOK NVFEVSKTODINRNINSHODDSEPOKOK KEPHROTTEVNIDSETVIRRALLSAFSENDIK VORNERNTAAEEHVAQODATLEHSTULDSSPS LSSTVVVPLRESYDPDVIPLENTLISTBLSSPS LSSTVVPLRESYDPDVIPLENTLISTBLSSPS LSSTVVPLRESYDPDVIPLENTLISTBLSSPS LSSTVVPLRESYDPDVIPLENTLISTBLSSPS LSSTVVVPLRESYDPDVIPLENTLISTBLSSPS LSSTVVVPLRESYDPDVIPLENTLISTBLOKSTA STSRADHSALPINGSLVRESSEVLKTSDSKEGG EGIFTVDTFAKASITSKRHIPEAHQATLLOGKOK OKNIMPLOSKLTGVIVENENTIKEGOLYOMA KKENDLANAENILKOTIKATVENOKKODGIAVD HVOGINTEKVAETVIKLKIKRSPKOVKDISID VORNENSEVUTSAGSGARSVLHORNOOTE DVATOPRRAEKTSVATSTEGOKKOVTLSPVK AGPATITSSETRQSEVALFOTSIEADEGLIUT THENNIPLHVGALASECTVFAAAEGGAVVTE GFAESETFILTSTKEGESGECAVAESEDRAADI LAVIHAVKIEANNINVTEKEDDAVTSAGSEE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE KCDGSLSRDSEIVEGTITFIESVESDGAVTSAGSE GESAVTSTOTTEDGEGGRASCTIGSEDSSEGRAMS SESEERGSAMDSTVAKSTOTVENVOHOTE GGSAVTSTOTTEDGEGGRASCTISTGCRESHINDT JDICSSKAGUPESSTVAKSGTOWNOTTE GESAVTSTOTTEDGEGGRASCTISTGCRESHINDT TOTTGGAEGGSDINFVICSVTGGGRREEMVTT GGSAVTSTOTTEDGEGGRASCTISTGCRESHINDT THERMIT THE TOTTGGAEGGEGOVTE FRAGAGGAMNANENNYDSMSGTEKGSKDT JDICSSKAGDVESVEROTTSTGCREGEDHTSVE NEECDGLMATTASGDITNONSLAGGKNOGK VIJISTSTNDYTPOGGCE GGMTSAASDQSDSQLEKVEDTTTSTGLVEGGS TVVLVSGEVPECEVALTSTSTEGEDHTSVERO NEECDGLMATTASGDITNONSLAGGKNOGK LISTSTNDYTPOGGCE GGPTSAASDQSDSQLEKVEDTTTSTGLVEGGS TOYLOVAGVEREETGEGFEPPTAGGSSTTATCASS LOPPAAAVEERGEGGBOPSVTATERGOTAMIST TSECCEAVMIGGLEGGABDSTITKOCO GOLSATEVSKHKVPAMSLAGRENGSTITADPEGON GO							KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
TERRISHIRNERLIS VICKDORCY SEVILITED NYRKENNIKKERRIS JASKIT KABHIS KRESSDSK 1QKDSLGSKQHOTTLORRSESYSEDKCOMDS) NDDSNLKPEEVVUKKERRITKSLLEEKLVLKS 1 KSKTOGKOVKVVETELGBGATKGATTERPD KEKNITENDSEK GORK SKVENDERPETGVEPV LETASSSAHSTOKDSSHRAKLPLAKEKYKSD KDSTSTELERKLSJOHKSSKLEHRSKDIKKKD ENKSDDKDGKEVDSSHEKARCINSLMEKKL BRILCENRIGSLSQEMAKGEEKLAANTLSTP SOSALGRPKKSGUDMILTPEGEPREIDSEPOVE NYFEVSKTOONRINNISHGIDISENIKGKD ATVOKDELRICTADSKATAA YPFORGTOV NSNSEKHADDHRSTLTKKMHIQSAVSKMNFGE EKPHRIGTTEVNIDSSTVHRILLSAFSENDER QKNLKNITAAEEHVAQGDATLEHSTNLDSSPS LSVYTVPLRESYPDVIPLFDKRITVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSEGG GEFTYTOTTPAKASTISKRHIPEAHGATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGL VDM KKENDLNAEPILKGTTKAVHLORGKOGAVD HVVGLNTEKVAETVKLKHKRSPGKVKDISID VERRINSEVOTTSAGGSAFSVLHORNOQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDERDATDL LAVHAVKIRANVSVYTEEKDDAVTSAGS KOLSSPSSIEVERVALPCTISTEGSEGCAVASEE KCDGSLSRDSEIVERSUTTFISEVESDGAVTSAG TERRAGSISSEEVDGSQONMIREMGPKERTER GASESTFTLTTKEGESGECAVASEEDRADL LAVHAVKIRANVSVYTEEKDDAVTSAGS TERRAGSISSEEVDGSQONMIREMGPKERTER GAGSVLGDDAPPGTASAGEGGASVDGTE GESAVTSTOTTEGGEGPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGE GESAVTSTOTTEGGEGPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG GENTSTAASDQSBSCLEVEDTTISTEVESCEGAL LAVHAVKIRANVSVTTEREDORSVINGTE GESAVTSTOTTEGGEPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG OTHER GASSESTERMENDLAGGRONG SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG GENTSTAASDQSBSCLEVEDTTISTIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC LYGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC LYGGS ATT STAGETGAASDCSABSTORMENOLUTATION TORT THE SEMBEGTER VITTEFFET TORT TORT TORT TORT TORT TORT TORT TO							KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
TERRISHIRNERLIS VICKDORCY SEVILITED NYRKENNIKKERRIS JASKIT KABHIS KRESSDSK 1QKDSLGSKQHOTTLORRSESYSEDKCOMDS) NDDSNLKPEEVVUKKERRITKSLLEEKLVLKS 1 KSKTOGKOVKVVETELGBGATKGATTERPD KEKNITENDSEK GORK SKVENDERPETGVEPV LETASSSAHSTOKDSSHRAKLPLAKEKYKSD KDSTSTELERKLSJOHKSSKLEHRSKDIKKKD ENKSDDKDGKEVDSSHEKARCINSLMEKKL BRILCENRIGSLSQEMAKGEEKLAANTLSTP SOSALGRPKKSGUDMILTPEGEPREIDSEPOVE NYFEVSKTOONRINNISHGIDISENIKGKD ATVOKDELRICTADSKATAA YPFORGTOV NSNSEKHADDHRSTLTKKMHIQSAVSKMNFGE EKPHRIGTTEVNIDSSTVHRILLSAFSENDER QKNLKNITAAEEHVAQGDATLEHSTNLDSSPS LSVYTVPLRESYPDVIPLFDKRITVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSEGG GEFTYTOTTPAKASTISKRHIPEAHGATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGL VDM KKENDLNAEPILKGTTKAVHLORGKOGAVD HVVGLNTEKVAETVKLKHKRSPGKVKDISID VERRINSEVOTTSAGGSAFSVLHORNOQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDKDVTLSPVK AGPATTTSSETTIGSEVALPCTISTEGKDERDATDL LAVHAVKIRANVSVYTEEKDDAVTSAGS KOLSSPSSIEVERVALPCTISTEGSEGCAVASEE KCDGSLSRDSEIVERSUTTFISEVESDGAVTSAG TERRAGSISSEEVDGSQONMIREMGPKERTER GASESTFTLTTKEGESGECAVASEEDRADL LAVHAVKIRANVSVYTEEKDDAVTSAGS TERRAGSISSEEVDGSQONMIREMGPKERTER GAGSVLGDDAPPGTASAGEGGASVDGTE GESAVTSTOTTEGGEGPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGE GESAVTSTOTTEGGEGPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG GENTSTAASDQSBSCLEVEDTTISTEVESCEGAL LAVHAVKIRANVSVTTEREDORSVINGTE GESAVTSTOTTEGGEPASCTGSSEGAMS SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG OTHER GASSESTERMENDLAGGRONG SESSEENGESSAMDSTVAKEGTIVPLVACAFGEG GENTSTAASDQSBSCLEVEDTTISTIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC VGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC LYGGS YDVLVSGVPECEVAHTSPSEKEDEDITTSVE MEECDGLMATTASGDTITATIC LYGGS ATT STAGETGAASDCSABSTORMENOLUTATION TORT THE SEMBEGTER VITTEFFET TORT TORT TORT TORT TORT TORT TORT TO	1						DETELHSSEKGLKVEENIOKOSOOTKLSSDDK
NVRKENNIKERRI.SAEKTKAEHKSRRSISSE (OKDSIGSKOHGTICRRSESYSEDECOMDIST NMDSNIL KPEEVVIKERRRTK SIL EERLY LIKE K SKENTOGKOVKVVETEL GEGATKOATTPKPD KEKNTEENDSEK ORKSVEDDE PFETTOVEPV LETASSSAASTORDSSHRAKLPLAKEKYKSD KOSTSTRIERKI.SIOGHSKSILISHSKIKKE ENKSDDIKDGEVDSSHERARCHSSONIKKKE ENKSDDIKDGEVDSSHERARCHSSONIKKKE ENKSDDIKDGEVDSSHERARCHSSONIKKKE ENKSDDIKDGEVDSSHERARCHSSONIKKKE ENKSDDIKDGEVDSSHERARCHSSONIKKEN SOSSLORPKKSGDMTLIPEOEPMEIDSEPOVE NYFEVSKTODNINNISHODIDSENMIKOKTS ATVOKDELRTCTADSKATAPAYKPGRGTOV NSNESKHADHRSTITKKMHIQSAYSKMNPGE KEPHRGTITEVNIDSETVHRMILSAPSENDRY OKNILINTAAEEHVAQODATLEHSTNILDSSPS LSSVTVVPLRESYDPDVIPLTDKRTVLEGSTA STSPADHSALPNGSLVTVRESVIKATSDSKEGG EGFTVDTPAKASITSKRHIPBAHQATLLOGKOK OKNIMPLOSKLTOVENDENTKEGKOGLOWA KKENDLNAEPMLKOTIKATVENOKKDGIGAVOV HVOGLITEKVAETVIKLHKRSPKOKVDISID VERRINESSEVDTSAGSGARSVLHORNOOTE DVATOPRASKITSVATSTECKOKDVTLSPVX AGPATTTSSETRGEVALPCTSHEADEGLIGT HSRNNPLHVGABASECTVFAAAEEGGAVVT GRAESETFLISTKEGESGECAVAESDRAADL LAVHAVKIRANVRSVTEIRKDDAVTSAGS EERLE KCDGSLSRDSEIVEGTTHISEVESGAVTSAG TEIRAGSISSEEVDGSQONMMEMOGYKKETEG GISAVTSTGTICHGAGRESDSKGTAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAREEGEGPAVSTSGOSSSGAAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAREEGEGPASCTGSSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAREEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD DEGITYSTGAKEEGEGPASCTGSDSSGGAJS SSSEENGESAMDSTVAKEGTNIPLVAAGPGCD OKSAKGIVESSYTSAVSGKDEVTPVFGCG GPMTSAASDOSSDSCLEVEDTTSTGLVGGG SYDVLYSGGVESSYTAAVSGKDEVTPVFGCG GPMTSAASDOSSDSCLEVEDTTSTGLVGGGS TOVLYSGENEEGEGPASCTGSGSTSGAGGS TOVLYSGENEEGEGPAGTGGGGSTSTSCOLO OKSAKGIVESSYTAAVSGKDEVTPVFGCG GPMTSAASDOSSDSCLEVEDTTSTGLVGGGS TOVLYSGENEEGEGPAGTGGGGSTSTANDT TSGCCAVMIGALADDIPONITATATAGGGST ASSAGGGLEGRANSPAHLRGPGGGGSSTSTSVOL	1						
IQKDSLGSKQHGITLQRESEYSEDKCDMDST NMDSNLKPEEVVYEKERRIKSLLEEKLVLKS KSKTOGKQVKVVETELQEGATKQATTREPD KERNTEENDSKQRKSKVEDRFEETGVEPV LETASSSAHSTQRDSSHRAKEPLAKEKYKSD KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD ENKSDDKDGKEVDSSHERAKEPLAKEKYKSD KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD SRRLCENRRGSLSQEMAKGEEKLAANTLSTT GORGEN NYFEVSKTQDNRNNISHQDDSEMMKQKTS ATVQKDELRTCTADSKATAPAYRFGRGTOV NSNSEKHADHRSTLTKKMHQSAVSKMNFG EXPHINGTEVNIDESTYPRMLLSAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSVTVYPLRESYPDVPHPLDKTATLEGSTA STSPADHSALPYQSLTVRESEVLKTSDSKEGG GOFTVTOTTPAKASTSKRHIPBALQTATLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNABPNLKQTIKATVENGKKDGIAVV HVGGLNTEKVAETVKLKHKRSPKGVKDISID VERNENSEVTSKAKSTSHENDALQTATLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNABPNLKQTIKATVENGKKDGIAVV HVGGLNTEKVAETVKLKHKRSPKGVKDISID VERNENSEVDTSAGSGSAFSVLHQRNQGTE DVATOPRRABKTSVATSTEGDKDAVTLG GFAESETFLTSTKEGESGGCAVAESEDRAADL LAVHAVKIBANVBSVTEKEDDAVTSAGSGE KCDGSLSRDSEVEGTTIFISEVESDGAVTSAG GFAESETFLTSTKEGESGGCAVAESEDRAADL LAVHAVKIBANVBSVTEKEDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFISEVESDGAVTSAG GFAESETFLTSTKEGESGGCAVAESEDRAADL GRANSSEEVEDGSQGNMMEMGFKKSTEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GGGVULGDNAPGTSASGGGOSVDJOTTE GESAATSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTTVYAAGFCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGESSGVLICESAEGDSQGTVMEH GESAATSTGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKEGTTVYPVGCG GFAESTFLTSTKEGESGGCAVAESDRAADL SESSERASSGSGESGESGESGSTEKEDTTSTGLVGGS YDVLVSGGVESSYTSAVSGKDEVTPVPGCG GFAMSTASSGEFUNGTSTSGLGGGS YDVLVSGGVESSYTSAVSGKDEVTPVPGCG GGEGAVTSTGARGPULSTADSTEGLUNGGS YDVLVSGGVESSYTSAVSGKDEVTPVPGCG GGEGAMSTTSIGGEFELPISSATTIKCAES ULTSTSTTNDYTPQVSAITDVEGGLSDALTTS RSEEKDECAMISTSIGGEFELPISSATTIKCAES ULTSTSTNDYTPQVSAITDVEGGLSDALTTS RSEEKDECAMISTSIGGEFELPISSATTIKCAES ULTSTSTTNDYTPQVSAITDVEGGLSDALTTS RSEEKDECAMISTSIGGEFELPISSATTIKCES ULTSTSTTNDYTPQVSAITDVEGGLSDALTTS RSEEKDECAMISTSIGGEFELPISSATTIKCAES ULTSTSTADEMPSALSRUPHENDLITATAVEDLSDA AIISTSTTADEMPSALSRUPHENDLITATAVEDLSDA AIISTSTTADEMPSALSRUPHENDLITATAVEDLSDA AIISTSTADEMPSALSRUPHSAGOGH PSAVCAEKEERHGKECFEIGPFAGGGSST ASYSAGGLEGNANSPA							
NMDSNLKPEVVIKEKRRTKSLLEEKLVLKS KSKTOGKOVKVVETELQGGATKATTPRD KEKNTERNDSEKORKSKVEDKPPETGYGEVP LETASSSAHSTOKDSSPHAKUPLAKEKYKSD KKENTERNDSEKORKSKVEDKPPETGYGEVP LETASSSAHSTOKDSSPHAKUPLAKEKYKSD KOSTSTPLERKLSDGHKSRSLKHSSKDIKKKUD ENKSDDKLDGKEVDSSHEKARGNSSLMEKKL SRRLCENRROSLSQBMAKGEEKLAANTLSTP SOSSLQRFKKSGDMTLIPEQEPMEIDSEPOVE NVFEVSRKVGDMTLIPEQEPMEIDSEPOVE NVFEVSRKVGDMTLIPEQEPMEIDSEPOVE NVFEVSRKVGDMTLIPEQEPMEIDSEPOVE NVFEVSRKVGDMTLIPEGPMEIDSEPOVE NVFEVSRKVGDMTLIPEGPMEIDSEPOVE NVFEVSRKVGDMTLIPEGPMEIDSEPOVE NNSESKHADHRSTLTKKMHIQSAVSKMNPGE KEPHIRGTTEVNIDSETVHRMLLSAPSENDRY QKNLKNTAAEHVAQDATLEHSTNLDSSPS LSSVTVVPLRESYPDPVIPLFDKRTVLEGSST STSPADHSALPNOSLTVRESEVLATSTBSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVVVENENTIKEGLOVANA KKENDLNAEPILKQTIKATVENOKKDGIAVD HVGKINTEKVAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGARSVLHORNOTOTE DVATOPRRAEKTSVATSTEGKOKGVINDID VERRNENSEVDTSAGSGARSVLHORNOTOTE DVATOPRRAEKTSVATSTEGKOKGVANDA KKENDLNAEPILKQTIKATVENOKKDGIAVD HVGKINTEKVAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGARSVLHORNOTOTE DVATOPRRAEKTSVATSTEGKOKGVANDAT HVGKATSTSTRABEGGAVTTE GFAESETTI.TSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEKDDAVTSAGSE KCDGSLSRDSEIVEGTTTFISEVSDGAVTSAG TEIRAGSISSEVDGSQOMMRIMOFKKETEG GFAESETTI.TSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEKDDAVTSAGSE KCDGSLSRDSEIVEGTTTFISEVSDGAVTSAG TEIRAGSISSEVDGSQOMMRIMOFKKETEG GESAVTSTGTEDEGFASCTGSEDSSEGRAIS SESSEENGESAMDSTVAKEGTNYPLVAAGPED DEGIVTSTGAKEGEGGAVSTFORGEIGH ASTCTGJGEESEGVLICESAGGDSQGTVVDF GGCG GESAVTSTGTEDEGFASCTGSEDSSEGRAIS SESSENGESAMDSTVAKEGTNYPLVAAGPED DEGIVTSTGAKEGEGGPASTGREGONEGH ASTCTGJGEESEGVLICESAGGDSQGTVVDT GGCG GPMTSAASDQSDSQLEKVEDTTSTGLVGGS YDVLVSGGVESSYTSAVSGREFTYPVGGCG GPMTSAASDQSDSQLEKVEDTTSTGLVGGS YDVLVSGGVESSYTSAVSGREFTYPVGGCG GPMTSAASDGSDSQLEKVEDTTSTGLVGGS YDVLVSGGVESSYTSAVSGREFTYPVGGCG GPMTSAASDGSDSQLEKVEDTTSTGLVGGS YDVLVSGGVESSTAASDGREFERDITSTGLVGGS YDVLVSGGVESGVANGREFTYPVGGCG GPMTSAASDGREFAGROTEGT ASTAGGARDAAVERGRAFT GVULSTADPEGD GGLATTVSKHKYMPSLLAENNCRCPGPVY GCKEPGPVLAVSTEREGGINGGSVHRSAGQGH PSAVCAEKEERIGGREFATGGAGSST ASYSAGGLEGNANSPAHLRGPEGTSGGTAK SSSKS		i '					
KSKTOGKOVKVVETELOGATKOATTPKPD KEKNTEENDSKORK SKVEDNFPETGVEPY LETASSSAHSTOKDSSHRAKI.PLAKEKYKSD KDSTSTRLERKLSDGHKSRSILKHSKOLKKKD ENKSDDKDOKEVDSSHEKARGNSSLMEKKL SRRLCENRGSLSQEMAKGEEKLAANTLSTP SGSSLQRKKSGDMTLIPEGEPMEIDSEPGVE NVFEVSKTQDNRNNNSKQDDSENMKOKTS ATVQKDELTICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMNFGE KEPHRGTTEVNIDSETVHRMLISAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNILDSSPS LSSVTVPLRESSYDPVPIPFDKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASTISKRHIPEGGL VDMA KKENDLNAEPHLKQTIKATVGATLLDGKO GKVIMPLGSKLTGVIVENENITKEGGL VDMA KKENDLNAEPHLKQTIKATVGATLLDGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVGATLLOGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVGATLLOGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVGATLLOGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVGATLLOGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVGATLLOGKO GKVIMPLGSKLTGVIVENENITKEGGLIVDMA KKENDLNAEPHLKQTIKATVAALEGGAVVTB DVATOFRAERETSVATSTEGGLOKADTLOGKE GFASETTLTSTKEGESGCAVAESEDRADL LAVHAVKLEANNSVYTEEKDATLSGGAVTSAG TEIRAGSISSEVEDVITSTGKAGGGEGAVTSAG TEIRAGSISSEVEDOSQGNMMRMOPKKETEG TVTCTGABGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPFGTSAGSGDGSVNDGTE GESAVTSTGITEDGEGAASCCTGSEDSSEGAIS SESEENGESAMDSTVAKGGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESGVLCESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMGTEKGSKDT DICSSAKGTVESSVTSAVSKOEVTPYPGGGE GFMSAASDQSDSQLEKVEDTITISTGLVGGS YDULVSGEVPECEVAHTSPSEKEDEDITISTYE NEECOGLMATTASGDTTINGNSLAGGKNQGK VLIISTSTTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFABMPSAVGROEDGLSVALRTE ENMEGTRVTTEEFFABMPSAVGROEDGLSVALRTE ENMEGTRVTTEEFFABMPSAVGROEDGLSVALRTE ENMEGTRVTTEEFFABMPSAVGROEDGLSTANDPEGN GILSATEVSKHKVPMFSLAERNNCRGFEVR GGKEPGVLANSTEEHNOPGFVRGFORP GGKEPGVLANSTEEHNOPSTHRAGGGTAK DSAVGERBERGANTASGLITTRVELDAA AIISTSTACCMPRASAIDRIERGFORPTGGGFG FANVASTEEHNOPSTHRAGGGTAK DSAVGEREKERAGECGFFGFRAGROGKESTI, HLINAEEKNVLLINSLOKEDSKSPTGTAGGSST ASYSAGRCLECNANSTAHLRGFEGTSGGTAK DSSVSIRVLANSTALLANDRIPGGTAK DSSVSIRVLANSTALLERGEGTSGATKA							
KEKNTEENDSKORKS, SVEDKPFEETGVEPV LETASSASAHTSORDSSHRAKIPLAKEKY YESD KDSTSTRIERK1.SDGHKSRSLKHSSKDIKKKD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRRLCENRRGSI.SQEMAKGEEKLANNTLSTP SGSSLQRKKSGDMTLIFEQEPMEIDSEPGYE NVFEVSKTQDNRNNSNSIQIDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTITKKMHDSSENMEW KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEEHVAGQDATULERSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGGTA STSPADHSALPNOSLTVRESEVLTSDKSEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGVVPLRENTTKEGGLVDMA KKENDLNAEPNIKGTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINSEVDTYSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGGATTTSSETRGSFVLHPGTSTEADEGLIIGT HSRNMPLHVGAEASECTVFAAAEEGGAVTE GFAESEFILTSTKEGGSGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKKDDAVTSAGSEE KCDGSLSRDSEIVGGTTTESPGDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVULGDNDAPFGTSAGSGEOSVNDGTE GESAYTSTGTEDGEGPASCTGSEDSSEGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESESVTSAVSGKDEVTYPVGGGC GRMTSAASDQSDSQLEKVEDTTTSTGLVGGSV YDVLVSGEVPCEVAHTSPSKEEDDITTSVE NEECDOLMATTASGDTNNNSLAGGKNQGK VLISTSTTINDYTTPQASATTOPLOGGLIATE EMMEGTRYTTEVEGASTOPLOTTEDLISSATTIKGASS (VCOYSENERAASDQSDSQLEKVEDTTTSTGLVGGS YDVLVSGEVPECEVAHTSPSKEEDDITTSVE NEECDOLMATTASGDTNNNSLAGGKNQGK VLISTSTTINDYTTPQAATDFEGRMPSAPP EAESPLASTISKEEKDECALISTISIAEECEASVS GVVVSENERAGTATVEDLADALARIE EMMEGTRYTTEVEGASTOPLARTE EMMEGTRYTTEVEGASTOPLATRE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEGASTORLARTE EMMEGTRYTTEVEDLADA AUISTSTAECKBEROPVTPAECEGNERTH ELEMBEGTRYTTEVEDLADA AUISTSTAECKBEROPVTPAECEGNERTH ELEMBEGTRYTTEVEDLADA AUISTSTAECKBEROPVTPAECEGNERTH HLINAEENVILLINSLYCEBESPYHTSAGCGH PSAVCAEKEEKHGKECPEIGFFAGROGKESTI. HLINAERNYLLINSLYCEBESPYHTSAGCHA BSSVASGRICLEGNANSFAHLRGFEGTSAGGAK BSSVASGRICLEGNANSFAHLRGFEGTSAGGAK BSSVASGRICLEGNANSFAHLRGFEGTSAGGAK							The state of the s
LETASSSAHSTQKDSKHRAKLPLAKEKYKSD KDSTSTRLERKLSOGHKRSKLSKOLKKKU ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRRJCENRRGSLSQEMAKGEERLAANTLSTP SGSSLQRPKKSGDMTLIPEGEPMEIDSEPGYE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDERTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMRYGE KEPHRGTTEVNIOSETVHRMLLSAPSENDRY QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVPLRESTYDPVPHEIDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIEGAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKGIGAVD HVVGLNTEKYAETVKLHKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNGOTE DVATGPRRAEKTSVATSTEGGLOVDMA KKENDLNAEPNLKQTIKATVENGKKGOTAV HVVGLNTEKYAETVKLHKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNGOTE DVATGPRRAEKTSVATSTEGGLOVDTAYA AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVQAAESECTVFAAAEEGGAVVTE GFAESETFLTSTKEGGSGCAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSADSEIVCGTTTIESEVSSGGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGABGRSDNPVICSYTGSESDSGGAIS GSSATSTGITBGGGPAACTGGE GESAVTSTGITBGGGPASCTGREEMYT GAGVVLGDNDAPPGTSASQEGDOSVNDGTE GESAVTSTGITBGGGGASORGEEMYT GAGVVLGDNDAPPGTSASQEGDOSVNDGTE GESAVTSTGITBGGGGASCREEMYT GAGVVLGDNDAPPGTSASQEGDOSVNDGTE GESAVTSTGITBGGGBAGTGEEMYT GAGVVLGDNDAPPGTSASQEGDOSVNDGTE GESAVTSTGTBGGGGSSSEGGAIS SSEENGESAMDSTVAKEGTNVPLVAAQPCD DEGIVTSTGAKEEDEEGEDVYTTSTGGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGTVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTTSTGLVGGS YDVLVSGEVVECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTINQNSLAGGKNOGK VLUISTSTTNOYTPQVSATIDVEGGSDATATIST ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATOPVLISTADEEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSESNERALADTRUKSCGGISTSSVEDC GGPVSSAAPQEEGDPSVTPAEEMGDTAMISTS TSEGGCAVMIGGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASITREEGNOCTADAPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVP GGKEPGPVLANSTEEGHNOPSVHCPSAQGH PSAVCAKEEKHRIKEFEGTTAGGSEST ASYSAGRCLEGNANSPAHLRGPEGTSCGTAK DSSVSGRVLAANTGAKADDMPPVQGTVA	1						, , , , , , , , , , , , , , , , , , , ,
KDSTSTLERK_DGHKSKLMSKLKHSKDIKKKD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRRLCENBRGGLSQEMAKGEEKLAANTLSTP SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE NVFEVSKTQDNRNNNSHQIDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSSKHADHRSTLTKKMHDSSENDRV QKNLKNTAAEHVAGGATVHRMLLSAPSENDRV QKNLKNTAAEHVAGGATVLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGGVEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGGVEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGGVEVLKTSDSKEGG EGFTVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGGLKDMTVSPVK AGPATTTSSTFRGSEVLHCTSIEADEGLLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVTE GFAESETFLTSTKEGESUGCAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTISEVSDGAVTSAG TEHAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPEERMVT GAGVVLGDNAPPGTSASQEGDGSVNDGTE GESAVTSTGTEDGEGASCTGSEDSSEGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEVTLSSAGEDSQLGTVVPL VAAGPCBESVTLSVSTGKEDSQLGTVVPLOGGE GPMTSAASDQSDSQLEKVEDTTSTGLGGSKDT DICSSAGGTVESSVTSAVSGKDEVPVPGGGC GPMTSAASDQSDSQLEKVEDTTSTGLYGGS YDVLVSGEVPECEVAHTSTSEKEDEDITTSVE NEECDGLMATTASGDTNQSLARGE WIESTTINDYTFQVSAITOPEGGLSDALRTE EMMEGTRYTTEEFAAMPSAVGDDSQLTAS RESEKDECAMISTSIGEEFELPISSATITIKCAES LQPVAAAVEERATOPLUSTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATTVEFLAADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATTVEFLAADFEGFWPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATTVEFLAADFEGFWPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATTVEFLAADFEGFWPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATTVEFLAADFEGFWPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSENERRATGTVMEKAGSGISTSVEDC GGKEGPVALAVSTEEMHORPSVHTSAGQGH PSAVCAEKEEKHGKECPEIGFFAGROGKESTI. HLINAEEKNYLLNSLKEEKDSPSVHTSAGGCH PSAVCAEKEEKHGKECPEIGFFAGROGKESTI. HLINAEEKNYLLNSLKEEKDSPSTGTAGGGSSI ASVSAGRCLEGNANSFAHLRGFEGTSAGGTAK DSSVSIRVLAAAVNTGAKADDMPPVQGTVA							· · · · · · · · · · · · · · · · · · ·
ENSEDDKDGKEVDSEHKARGNISLMEKKL SRRLCENRRGSLSQEMAKGEKLAANTLSTP SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELKTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSKMPGE KEPHRGTTEVNIDSETYPHRMLLSAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSYTVVPLRESYPDVPLYDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENSHIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKGGRAVD HVVGLNTEKYAETVKLKHRSPGKVKDISID VERNINSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAKTSVATSTEGKDKDVTLSPVK AGPATITTSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVQAEASECTVAAEEGGAVUTE GFASSETFLTSTKZGESGECAVAESEDRADJL LAVHAVKLEANNSVYTEKDDDAYTSAGSEE KCDGLSRDSEIVEGTTITISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRNGPKETEG TVTCTGAEGRSDNFVICSTGEDSSEGFALS SESEENGESAMDSTVAKEGTNVPLVAAQPCD DEGIVTSTGAKEDEGEGFATS SESEENGESAMDSTVAKEGTNVPLVAAQPCD DEGIVTSTGAKEDEGEGFATSTGEGSGGASTONDGTE GESAVTSTGTIEDGGEPSATVSTGGRPEERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTIEDGGEPSATSTGEGSGGTAVESCHOLD SESEENGESAMDSTVAKEGTNVPLVAAQPCD DEGIVTSTGAKEDEGEGGTAVSTGEGRSGGTATS SESEENGESAMDSTVAKEGTNVPLVAAQPCD DEGIVTSTGAKEDEGEGGTATSTGERGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTSUTSVE NECCOGLMATTASGDITHQNSLAGGKNQGK VLILISTSTTINDTYPQVSATIPSGERGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTSVE NECCOGLMATTASGDITHQNSLAGGKNQGK VLILISTSTTINDTYPQVSATIPSGERGSKDT SEEKDECAMISTSIGEEFELPISSATITKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAP EAESPLASTSKEEKDECALISTISAGECASSYS GVVVESENERAGTVTMEERGGSGSISTSSYEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSSEGCAVMIGAVLQDBERNORCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKYPSAGQGH PSAVCAEKEEKHIGKECFEIGPFAGRQKESTI. HLINAEENVALLINSLQKENNORCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKYPSAGQGH PSAVCAEKEEKHIGKECFEIGPFAGRGKSTI. HLINAEENVALLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGMANSPAHLROPEGTGGTAK DSSVSSIKYLLAAVNTGGLARDDMPPVQGTVA	1						1
SRRLCENRRGSLSQERMAGGERLAANTL.STP SGSSLQRPKKSGDMTL.PEOEPMEIDSEPGYE NVFEVSKTQDNRNNISHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGGTGV NSNSEKHADHRISTLTKKMHIQSAVSKMRPGE KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEHVAQGDATLEHISTINLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPNQSLTVYRESVLKTSDSKEGG EGFTVDTTAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNABEPNLKQTIKATVENGKKDGJAVDH HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINENSEVDTSAGGSASVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPI.HVQAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKLEANNVSVVTEGESTAA LAVHAVKLEANNVSVVTEGENDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGHTEDGEOPASCTOSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGITVSTGAKEEDEEGEDVVTSTGRGNEICH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDTTPQVSATTDVPGGGLSDALRTE ENNEGTRVTTEEFEAPMPSAVSGDDSQLTAS RESEKDECAMISTSIGGEPELPIISSATTIKCAES LQPVAAAVERAATGPVLISTADPEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECECASVS GVVVESENERAGTVMEERGGSGISTSSVEDC EGFVSSAVPQEEGDPSVTTAEERGDSQLATAS RSSEKDECAMISTSIGGEFELPIISSATTIKCAES LQPVAAAVERAATGPVLISTADPEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECECASVS GVVVESENERAGTVMEERGGSGISTSSVEDC EGFVSSAVPQEEGDPSVTTAEERGDSQLATAS RSSEKDECAMISTSIGGEFELPIISSATTIKCAES LQPVAAAVERAATGPVLISTADPEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECECASVS GVVVESENERAGTVMEERGGGSGISTSSVEDC EGFVSSAVPQEEGDPSVTTAEERGDSGLATAS TTSEGGCAVMGGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGO GDLSATEVSKIHV VPMPSJALAENNCPCPGPVR GGKEPGPVAVASTEEGHNGPSVHKPSAGQGH PSAVCAEKEERHGKSCPEIGPFAGGSST ASYSAGRGLEGMANSPAHLROPEGTGGGTXK BURGERGGATAANSPAHLROPEGTSGGTAK DSSVSSIKYLAANVTIGAKADDMPPVQGTVA	1						
SGSSLQRPKKSGDMTLIPEQEPMELDSEPGVE NVPEVSKTODNRNINSHQDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSA YSKMYPGE KEPHRGTTEVNIDSETYHRMLISAPSENDRV QKNILKNTAAEEHVAQGDATLEHSTNIDSSEPS LSSVTVVPLRESYPDPVIPEDKRTVLEGSTA STSPADHSALPNQSLTVRESPULTSDSKEGG GEGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATKHKRSPGKVKDISID VERNINSSEVDITSAGSGSAPSVLHQNRQGTE DVATGPRRAKTSVATSTEGKDKDVTLSPVK AGPATITTSETRQSEVALPCTSIEADGLIIGT HSRNNPLHVQAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIRANVASVYTEDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRNGPKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGJGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKAGGTNVPLVAAGPCD DEGIVTSTGAKEDEGEGGNTSTEGSKENDT GAGSVLGDNDAPPGTSASQEGJGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKAGGTNVPLVAAGPCD DEGIVTSTGAKEDEGEGGNVTSTGRGROEGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGTVSESVTSAVSGLOEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTISVE NEECDGLMATTASGDITNQNSLAGGKNOGK VLIISTSTTINDTYPQVSATGVEGGLISTSVE NEECDGLMATTASGDITNQNSLAGGKNOGK VLUISTSTTINDTYPQVSATGVEGGSLGALARTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGGEPELPISSATTIKCAES LQPVAAAVEERATGPVILISTADFEGFMPSAP EAESPLASTSKEEKDECALISTSJAECEASSV GVVVESENERAGTVTMEERGGSGISTSSVEDC EGGVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDTTITTVELSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHVPMPSATGRAMSTETAGSST ASYSAGRGLEGMANSPAHLRGPEGTSGGTXA BYSAGRGLEGMANSPAHLRGPEGTSGGTTAK BUSSVSIKYLAANVIGANDAPPVQGTVA	1						
NVFEVŠKTODNRNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKA PAYKPGRGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPIHRGTTEVNIDSETVHRMLSAPSENDRV QKNLKNTAAEHVAQGDTLHSITNLDSSPS LSSVTVVPLRESYDPDVIPLFDKRTVLGSTA STSPADHSALPNQSLTVRESVLKTSDSKEGG EGFTVOTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KENDLNAEPNLKQTIKATVENGKKDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD HVVGLNTEKY AETVALKHRRSPGKVDGJAVD GRAPATITSSETTQSEVALPSIEADEGLIJGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFTISEVESDGAVTSAG TEIRAGSISSESEVDGSQGNMRMMMPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSAGGGGGSVNDDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDOSDOLEKVETTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGNNQGK VLIISTSTTNDVTPQVSATITDYPGGGESDALRTE ENMEGTRYTTEEFEAPMPSAVSGDDSQLTAS RESEKDECAMISTISGEFEISSATTIKCAES LQPVAAAVERAATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERACTVMEERGOGGISTSSVEDC EGFVSSAVPQEEGDPSVTTAEEMGDTAMISTS TSEGCCAVMIGAVLQDEDTTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTS TSEGCCAVMIGAVLQDEDTTRVEUSDA AIISTSTAECMPISASIDRHEENQLTAMISTERGORST ASYSAGRGLEGMANSPAHLROPECTSGGTAK DSSVSSIRYLAANVTGAANSPAHLROPECTSGGTAK DSSVSSIRYLAANVTGAANDAPPVQGTVA	1						l .
ATVQKDELRTCTADSKATAPAYRGRGTGV NSNSEKHADHRSTLTKMIQSAVSKMMPGE KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEEHVAQQDATLEHSTNLDSSPS LSSVTVVYPLRESYDPDYPLFDKRTVLGGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GEGFVTOTPAKASITSKHPEAHQATLLDGKQ GKVIMPLGSKL TGVIVENENTIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKIKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFILTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSSRDSEIVEGTTTIFISEVESDGAVTSAG TEIRAGSISSEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNPVCVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSBDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTYSTGRGNEIGH ASTCTGLGESSGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKDEVTPVPGGCE GFMTSAASDQSBOSLEKVEDTTISTGRVSEIGH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKDEVTPVPGGCE GFMTSAASDQSBOSLEKVEDTTISTGRVSEIG VULVSGEVPECEVAHTSPSEKDEDIITSVE NEECDGLMATTASGDTINQNSLAGGRNOGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS SESEENGECAMISTSTUGEFFELPISSATTIKCAES LQPVAAAVEERATGPVULSTADFGPPMPSAPP EASPLASTISREKDEGCALISTSIAEECEASVS GVVVESENERAGTVMEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQBEDRITTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GJUSTSTRAEMPSLAERNSTSSSESOED GRASSARSVENSTSARSKERDEGTAMSTST TSEGCEAVMIGAVLQBEDRITTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSETITAGGSST ASYSAGRGLEGNANSPAHRGPGFTSGQTAK BUSSNSSIRVLAAVNTGAKADDMPPVQGTVA							
NSNSEKHADHRSTLTKKMHIQSA VSKMMPGE KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNIKNTAAEEHVAQDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNQSLTRSEVLKTTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGJAVD HVVGLNTEKYAETVALKHKRSPGKVKDISID VERRIENSEVDTSAGSAPSVLHQRNGOTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERNVT GAGVVLGDNDAPPGTSASGEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAIMMANENNVDSMGSTEKGSKDT) DICSSAKGIVESSVTSAVSKDEVTPVPGGGE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVECEVAHTSPSEKEDDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGITNTTEFFEAPMPSAVSGDDSOLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERAATGPVLISTADFEGPMPSAPP EASPLASTSKEEKDECALISTSIGAEECEASVS GVVVESENERAGTVMEEKDGGSISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHVKPWETEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHVKPUSTEGMISTSSVEDC EGPVSSAVPQEEGDPSVTREEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGINGPSVHKPSAGQGCH PSAVCAEKEKHGKECPEIGPPAGERGOKESTI. HLINAEEKVLLINSLQKEDKSPETIGTAGGSST ASYSAGRGLEGNANSPAHLRGPETSGQTAK	j						
KEPHRGTTEVNIDSETVHRMÎLSAPSENDRV QKNILKNTAAEHVAQGDATLEHSTNILDSSPS LSSVTVVPLRESYDPVPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLIKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNILKQTIKATVENGKKDGJAVDH HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGAPSVLHQRNGQTE DVATOPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASEGAPSVLHQRNGQTE GFAESETFLTSTKEGESGCAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSSDSEIVEGTTIFISEVESDGAVTSAG TEIRAGSISSEVUGSQGNMMINGPKKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTTEDGEGPASCTGSBDSSEGPASI SESEENGESAMDSTVAKEGTNVPLVAAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGESSGOVLICESAEGDSQIGTVVEH VEAEAGAAIMHANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDTNONSLAGGKNOGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVULSTADFGPPMPSAPP EASPLASTSKEEKDECALISTSIAFECEASVS GVVVESENERAGTVMEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEMGDTAMISTS TSEGCEA VMIGAVLQDEDRITTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATTSVSKHV PROBEDITANTSTS TSEGCEA VMIGAVLQDEDRITTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATTSVSKHV PROBEDITANTSTS TSEGCEA VMIGAVLQDEDRITTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANDSAHARGPCTSGQTAK							
QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTFAKASITSKRIJEFAHQATLLDGKQ GKVIMPLGSKLTGVLENENTIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGJAVD HVVGLNTERYAETVKLKHKRSPGKVKDISID VERRNESEVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVLPCTSSEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVSDGAVTSAG TEIRAGSISSEEVDGSOGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPRGTSASQEDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEDEEGEDVYTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLEARTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLARTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLARTE ENMEGTRVTTEEFEAPMPSAVSGDSQLARTE ENMEGTRATTERATERATIVATE							
LSSVTVYPLRESYPDRVIPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGITVDTTAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNILKTSVKHKAYENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHORNGOTE DVATGFRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGSGECAVAESEDRAADL LAVHAVKLEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVPGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGGEFELPISSATTIKCAES LQPVAAVEERATGPVLISTADEFGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGGEAVMIGAVLQDEDRLTTTREDLSDA AIISTSTAECMPISSATTIKCAES LQPVAAVEERATGPVLISTADFEGNPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDSGIISTSSVEDC EGPSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGGEAVMIGAVLQDEDRLTTTREDLSDA AIISTSTAECMPISSATTIKCAES LQPVAAVEERATGPVLISTADFEGNPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDSGIISTSSVEDC EGPSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGGEAVMIGAVLQDEDRLTTTREDLSDA AIISTSTAECMPISSAIDRHEENQLTADNPEGN GDLSATEVSKHK VPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNORSVHKPSAGQGH PSAVCAEKEEKHIGKECPEIGFFAGRQKESTI. HLINAEEKNYLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSTRYLAAVNIGLKALKADDMPVQGTVA	1	i					KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLOSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVGLUTTEKYAETVKLKHKRSPGK VKDISID VERRIENSEVDTSAGSGSAPSVLHQRNGQTE DVATOPRAEKTSVATSTEGKKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGSGGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSADSEIVEGTTITISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMOFKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEEGDGSVNDGTE GESAVTSTIGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDGEDVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDVTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGGEFELPISSATTIKCAES LQPVAAAVERATGFYDVSAGTBACKES GVVVESENERAGTVTMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDTAMSTTSTSCECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDTAMSTTS STEGGEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECHISASIDRHEEROLTADNPEGN GDLSATEVSKHK VPMPSLLAENNCEPGVR GGKEPGPVLAVSTEEGHNORSVHKPSAGQGH PSAVCAKEEEHIGKECPEIGPFAGRGKESTI HINAEEKNYLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSTRYLAAVVIRSLAERAGRAGKADSDPVQGTVA	1						QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPHLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLIKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHORNQOTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNIPLHVGAEASECTVFAAAEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQOIMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEOPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESEBNERAGTVMEEKDOSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEPMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGINGPSVHKPSAGGGH PSAVCAEKEEKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGINGPSVHKPSAGGGTAK DSSVSSIRYLAAVTGAIKADMPPVQCTVA)						LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVTE GFAESETFLISTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQQOMMRMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTTGIGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE EMMEGTRVTTEEFEAPMPSAVSGDDSQLAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQEBORLTTITTVEDLSDA AIISTSTAECMPISSAIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGINGPSVHKPSAGGCH PSAVCAEKEEKHGCPGIPFFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTFSQCTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1						STSPADHSALPNQSLTVRESEVLKTSDSKEGG
KKENDLNAEPPILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFILTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVISVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOFTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTTSTGRGNEIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VEAEAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLYGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDTYQVSAITDVEGGLSDALRTE ENMEGITRYTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVYMEEKDGSGIISTSSVEDC EGGPSSAVPQEGOPSVTTAEEMGDTAMISTS TSEGCEAVMGGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISAIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAKKEERHGKCPEIGPFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSCGTAK DSSVSSIRYLAAAVNIGAIKADDMPPQQGTVA	1				_ 1		EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
KKENDLNAEPPILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFILTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVISVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOFTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTTSTGRGNEIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VEAEAGAAIMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLYGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDTYQVSAITDVEGGLSDALRTE ENMEGITRYTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVYMEEKDGSGIISTSSVEDC EGGPSSAVPQEGOPSVTTAEEMGDTAMISTS TSEGCEAVMGGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISAIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAKKEERHGKCPEIGPFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSCGTAK DSSVSSIRYLAAAVNIGAIKADDMPPQQGTVA		•					
HVVGLNTEKYAETÜKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETITTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITPISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPRERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGFASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGGGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMTSTIGEFEFLISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVSESNERGTVMEKKGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADIPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGINGFSVTHRYSAGQGH PSAVCAKEKEERHGKECPEIGPFAGRGGKESTL HLINAEEKNVLLNSLOKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSIRYLAAAVNIGAIKADDMPPVQGTVA	ł l						l I
VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDUTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGCCAVAESEDRAADL LAVHAVKIEANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAJMANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIBSTSVEDC EGPVSSAVPQEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTACCMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQCH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLOKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQCTAK DSSVSSIRYLAAVNIGAIKADDMPPVQGTVA	1						`
DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECECASVS GVVVESENERAGTVMEEKOGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEKHGKECPEIGFFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1				19		1
AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETTLSTKEGESGECAVESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGISSEVDGSQGMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALSTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEKHKECPEIGFFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGQTAK DSSVSSIRTJAAVNTGAIKADDMPPVQGTVA	1						
HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASQDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERRATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMGAVLQDEDRLTITEVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKIKK VPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTIGAIKADDMPVQGTVA	1						1
GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EASSPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGPPSVTTAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNITGAIKADDMPPVQGTVA					Ì		-
LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITIFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROBEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							
KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPCPVR GGKEPOPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1				!		
TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHK VPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCCPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTIGAIKADDMPPVQGTVA	1						
TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENNMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHK VPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							1
GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRAEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECECASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					ļ		
GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDGKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTIGAIKADDMPPVQGTVA	1 :						
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1				i		,
DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNYDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			' I		l	1	
ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1			1	i		
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					.		
DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					. !]	
GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA]	l		
YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							
NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					ļ		
VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			[•	ĺ		
ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA						ļ .	
ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			ł	1	Į.		VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGVKESTL HLINAEEKNVLLN'SLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		· [1			ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLLN'SLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			j	}			RSEEKDECAMISTSIGEEFELPISSATTIKCAES
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLLN'SLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA) 1			ļ			LQPVAAAVEERATGPVLISTADFEGPMPSAPP
GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							•
EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	(ľ	ì		į į	
TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA				}	ł		
AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1			- 1		l	•
GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1					ļ	7
GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1				į		
PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	j !			ſ	ļ	ĺ	
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					l		
ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA				1	l		
DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1		- 1	ł	1	,	-
				Į	,]	, ,
EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP)		ļ	j	j	ļ	
	L				l		EHSPLPAEQQGSEDNLKTSTTKCITGQESKIAP

Section Sect	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide		, ,					
Sequence			,,,,,,				
Sequence							
1914 ng no first amino neitd residue of peptide pe	1						
amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence per solo codon, peposible nucleotide deletion, l-possible nucleotide deletion l-possible nucleotide deletion l-possible nucleotide deletion l-possible nucleotide deletion l-possible nucleotide deletion letion		uchico	1				
Peptide	uciico	•	1	/17			
Peptide Pept	İ						
	i	i	ĺ			sequence	
SITIMIPPATYSVALLAPKCEQDLTIKNDYSGK WTOQASAEKTQDINSTRKSPREGDIMVTVS SEEMVCDIQNESPLINULGQLKLKANILKMEA YVPSEEKNGEILAPPSLLGCKPREGAELQRE PLLVWESLNVENSGFRTNEEHISESYNKGEISS GRKDNAEAISGHSVEADPKEVEEEERHMEKR KNKQIYILSEDEPIDDNEDVLDSRIETAGRQC PTETPHATEENSRULEELKKTSSETISTISTS WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKKTSSETISTISTS WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKKTSSETISTISTS WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKTSSETISTISTST WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKTSSETISTISTST WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKTSSETISTISTSTS WEEKDEYSSETTGEKPEQNODDTIKSQE PTETPHATEENSRULEELKTSSETISTISTSTS WEEKDEYSSSETTGEKPEQNODDTIKSQE PTETPHATEENSRUPPRPPAFFSPELAGFEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPELLY VIPPKELAK YEYMERGEPH PTYTPLALADILE PTYTELLY VIPPKELAK YEYMERGENSY PTYTELLY VIPPKELAK YEYMERGEPH PTYTPLALADILE PTYTELLY VIPPKELAK YEYMERGEPH PTYTPLALADILE PTYTELLY VIPPKELAK YEYMERGEN PTYTELLY VIPPKEL PTYTELLY VIPPKELAK YEYMERGEN PTYTELLY VIPPKELAK YEYMERGEN PTYTELLY VIPPKELAK YEYMERGEN PTYTELLY VIPPKELOK YEYMERGEN PTYTELLY VIPPKELOK YEYMERGEN PTYTELLY VIPPKELOK YEYMERGEN PTY	1	ļ	ļ				
WITQASAEKTGDDNSTKKSFPEGDIMVITVS				 	Sequence		
]		ŀ				
1302 2652 A 10167 321 842 EFSLEPFLEVHYERINALESSINENGERSGARGALAGERSYNKGERS GRKDNAGALSGHSVEADPKEVEEERHMEKKRIKRIVHLSSEDEPDDMPDVLDSIKTAORQC PETEPHATKEENSRDLEELPKTSSETNSTTSTV MEKDEYSSETTGKEPQNADDIKSQB EFSLEPFLRFSFARPPREPAPFPSPELAGFEPH FVFYFFLSVHPPKELAKVEYMEEQVILTEKEN NSTVARGGTSVRCLSFSPRILPFLLLALLEL DGFGEHPFYHCLVAEVPKEENTTEGRVESPH FVFYFFLSVHPPKELAKVEYMEEQVILTEKEN SNSTVARGGTSVRCLSFSPRILPFLLLALLEL DGFGEHPFYHCLVAEVPKEENTTEGRVESPH FVFYFFLSVHPPKELAKVEYMEEQVILTEKEN SNSTVARGGTSVRCLSFSPRILPFLLLALLEL DGFGEHPFYHCLVAEVPKEENTTEGRVESPH EAREKCYVKSSVGCVEPLTTQAEVTENLDR KNSQQVFKLLKK SNSQQVFKLKK SNSQQVFKLKK SNSQQVFKLK SNSQQVFKLK SNSQQVFKLK SNSQQVFKLK SNSQQVFKLK SNSQQVFKLK SNSQQVFKLK SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFKL SNSQQVFK S	ļ	}	Ì		ļ		
PILLYNESLNYENGGRTNEEHISESYNKGERG RICHARCH RICHARC	i		Ì		1		
	1	l			İ		
	ĺ	1	1		[•
PETEPHATKEENSRÜLEELEKTSSENTISTERV	l	l	ŀ	İ		i	· ·
MERKDEYSSETTIGER/EQNDDTITISQGE			ł	ļ	ì		
1302					1		
	1202	2652	<u> </u>	10167	221	942	
NSTYAGRGTSVRCLSPSRPLPHLPLLADILE DGFGEHPFYELCLASPYFELTAGENTPSPPP	1302	2032	^	10107	321	042	
DGFGEHFYHCLVAEVPKEHWTPEGNPSPPP			!				
EARTKCYVRSSYGCVEPLTTQAEVTENLDR	l	1			1	l	1
NSSQOYFKILKKK NMILKERRILINSLGEGTINGLLDELLETINV LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITY LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITY LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF YHEAVVLFTQALKLNPQDHRLFGRRSFCHER LGQPAWALDAQWALTILRPGWPRGIFFRLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLIHLTLQGQRGGICAPPLSPGALQPSILL RSCLIHLTLQGQRGGICAPPLSPGALQPSILL RSCLIHLTLQGQRGGICAPPLSPGALQPSILL RSCLIHLTLQGQRGGICAPPLSPGALQPSILL RSCLIHLTLQGRGGICAPPLS	1	1					
1303 2653 A 10171 206 429		1					•
1304 2654 A 10184 970 1524 LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF	1202	2652		10171	206	420	
1304 2654 A 10184 970 1524 LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF	1303	2033	Α .	101/1	206	429	
1304	l	l	l	ļ	ŀ	ļ	
YHEAVVLFTQALKLNPQDHRLFGNRSFCHER LGQPAWALDAQVALTLRPGWPRGLFLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH 1305 2655 A 10194 2 394 TIDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPHR DFISVTLSFGES YDDSKSWRRSGWRKWKQL SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT 1306 2656 A 10195 1 410 PGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFITIVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFK WHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFK GGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR 1308 2658 A 10214 2 453 EGGIRQF0FGPPPALASAPAATMRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERINFLVEEYSTSGRL DNITQYMSLHTQYLSFIRSQFYMRMDGPL PLPYRHYIAIMAAARHQCSYLINM SRKHLCGGSILHWWWVLTAAHCFRTILDM AV AFQLFNAKCESAFLSKRNFLQRIWRMDGPL PLPYRHYIAIMAAARHQCSYLINM SRKHLCGGSILHWWWVLTAAHCFRTILDM AV APQLFNAKCESAFLSKRNFLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRNITGASLADI AV APQLFNAKCESAFLSKRNFLQRNWVVDDHJ DRDCDSHPGKKKEKKJETYTRCSKEGGCKKKEML QMVCAQCHGNFCIQHRPHDHSCRHGSRPTI KAG STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1004	0664		10104		1504	
LGQPAWALADAQVALTILRPGWPRGLFRLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGGGGGCAPPLSFGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH T01LGRFRFWGAMAACEGRFSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRMILFLLAFLLFCGLLFYNLADHWKG RNTCT 1306 2656	1304	2654	A	10184	970	1524	
ALMGLQRFREAAAVFOETLRGGSOPDAAREL RSCLIHLTLQQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLIH 1305 2655 A 10194 2 394 TDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRINMLFLLAFLLFCGLLFYINLADHWKG IRNTCT 1306 2656 A 10195 1 410 IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRGCVRLRGAVI GIDDEDDSTFITIVDQKTHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFFWLVSGFFFK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK SAFIPKEVQANTVKSR SPDQQNYTKSR 1308 2658 A 10214 2 453 ECGGIRQPGFPPPALASAPAATMNRVGGSPS SAAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPKEVVQANTVDERTINFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM RGWPEQQSTGRRDVARQPRCQKEEGRRLRP RALESRTFGGSERSWGPPLESTKENVQCGH RRAFPNSSWLPFHERLQVQNOECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRPLQRNWTVLYRRK HKKGQSAEIQKKKTTRRAFKFQRAITGASLADI MAK AFQLFNACCEGAFCKKKPML QMVCAQCHGNFCIQHRPLDHSCRHGSRFTI KAG STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			1				
RSCLIHIT.T.GGQRGGICAPPLSPGALQPLPHA	1			}		ľ	
ELAPSGLPSLRCPRSTALRSPGLSPLIH 1305	1	l	ĺ		f		
1305		1			ŀ	i	
SDFLTPPVGGAPWAVATTVVMYPPPPPPHR DFISVTLSFGESYDNSKWRRRSCWRK WKQL SRLQRNMILFLLAFLFCGLLFYINLADHWKG IRNTCT IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFITITVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SAFIPEKEVVQANITVDERTINFLVEEYSTSGRL DNTTQVMSLHTQVLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM SAFIPEKEVVQANITVDERTINFLVEEYSTSGRL DNTTQVMSLHTQVLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM RAJESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRRHLCGGSILHWWWVLTAAHCPRRTLLDM AV AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HCGGSHPGCHCSEKTCKQ LDFLPVKCDACQCPCKDHFFYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKQPDVVVGDHI DRDCDSHPGKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		-	ļ.,	1000			
DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRNMILFILAFILFCGLLFYINLADHWKG IRNTCT	1305	2655	A	10194	2	394	
SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT 1306	}	1	j		ļ	ļ	
IRNTCT	J]	1				
1306	ļ						
NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTITITVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK 1307 2657 A 10205 85 308 QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR 1308 2658 A 10214 2 453 ECGGIRQPGPPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKK GQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPVAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDH DRDCDSHPGKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1001						4
GIDDEDDSTFITIVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK	1306	2656	A	10195	ŀ	410	
WIHALEETILRHTLQLQVRVFTWFPDSSLVGA		ļ					
1307 2657 A 10205 85 308 QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR 1308 2658 A 10214 2 453 ECGGIRQPGPPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFFNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VITLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	l	l	ł				
1307 2657 A 10205 85 308 QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR 1308 2658 A 10214 2 453 ECGGIRQFGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDEERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSIRLFHLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	ì			ŀ		
DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR 1308 2658 A 10214 2 453 ECGGIRQFGPPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDEERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGPRDVARQPRCQKEEGRRLRP RALESRTTQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		4	ļ				
SPDQQNYTKSR	1307	2657	A	10205	85	308	
1308 2658 A 10214 2 453 ECGGIRQPGPPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTINFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKQQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	j	}]	1			
AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			 -	1			
SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1308	2658	A	10214	2	453	ECGGIRQPGPGPPPALASAPAATMNRVGGSPS
DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGPRDVARQPRCQKEEGRRLRP RALESRITFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	l	ľ	ł				
PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	}	1				
1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	ŀ	ł	l		1	!	
RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			L				
RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1309	2659	A	10233	45	421	
SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	j)	l				
AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHL DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	1	1				
1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP							1
HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP				L			
MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1310	2660	A	10241	243	442	, ,
1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	}	l	ł	1			
VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	L	<u> </u>	<u></u>	L	L	L	
LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1311	2661	A	10261	751	176	
QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1						
QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	l		l				LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	([1	[-			
QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1]	1	[
1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	ł	l	1		1	
	L	-	1				KAG
SMTILDKKDGEQAKALFEKVRKFRAHVEDSD	1312	2662	A	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP
	L	1	l	1			SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF EHVCKPKVEHLIGYEVFECTHNMAYMLKKL
						LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	A	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRGRSRSYSRSRSWSKERLRERDRD RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN LPGMQPFPAQPPVVEGPPPPGLPPPPILTPPPV NLRPPVPPPGPLPPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVYYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	Α	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NEKLVDEKTILETSFHQHRERAEQLSQENEKL MNLLQERVKNEEPTTQEGKIELEQKCTGILE QGRFEREKLLNIQQQLTCSLRKVEEENQGAL EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1318	2668	A	10303	333	879	EECRVTLEGLKMENGSLKSHLQG GECFIMAAVVQQNDLVFEFASNVMEDERQL GDPAIFPAVIVEHVPGADILNSYAGLACVEEP NDMITESSLDVAEEEIIDDDDDDITLTVEASCH DGDETIETIEAAEALLNMDSPGPMLDEKRINN NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL LIGAKSLPASVVLEAFSGTCQSADCTIVI.DAR LPRTLAGLLAGGALGLAGALMQTLTRNPLAD PGLLGVNAGASFAIVLGAALFGYSSAQEQLA MAFAGALVASLIVAFTGSQGGGQLSPVRLTL AGVXL
1320	2670	A	10323	441		KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV AVVDIQSDKAANVAQEINAEYGESMAYGFG ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI AKAAFISDFQLGDFDRSLQVNLVGYFLCARE FSRLMIRDGIQGRIIQINSKSDE
1321	2671	A	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD ILKCTLLVFGVRILYILKLNYTTEECDMKNMH YVDPDHVKRAQKYAQQVLQKESPPKFAKTS MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS ERKMRAHQVLTFLLLFVITSGASENASTSRGC GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL LGTLG
1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH QVLTFLLLFVITSVASENASTSRGCGLDLLPQ YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	Α,	10336	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE NSVTHHEVKCQGKPLAGIYRKREEKRNAGN AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE LQSEERKRIDELIESGKEEGMKIDLIDGKGRG VIATKQFSRGDFVVEYHGDLIEITDAKKREAL YAQDPSTGCYMYYFQYLSKTYCVDATRETN RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL RGVTATFGRPAEWPGYLSHLCGRSAAMDLG PMRKSYRGDREAFEETHLTSLDPVKQFAAWF EEAVQCPDIGEANAMCLATCTRDGKPSARML LLKGFGKDGFRFFTNFESRKGKELDSNPFASL VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE QLYQDQEVPKPKSWGGYVLYPQVMEFWQG QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE DWLYERLAP
1326	2676	Α	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1104	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ		ŀ	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}		}	j	peptide	ļ	/=possible nucleotide deletion, \=possible
			 	sequence	ļ	nucleotide insertion HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD
]	1		j	EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY
	1		ļ	Ì		LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP
]	1		}	NKKSPPEPRVAKKLGMIAGGTGITPMLOLIRA
	}					ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ
			1		}	ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD
}	i	Ì	į		İ	MIREHLPAPGDDVLVLLCGPPPMVQLACHPN
			<u> </u>			LDKLGYSQKMRFTY
1327	2677	Α	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
]				RYHRLSDMSMLAERRRKQKWAVDPQNTAW
	Ì		ļ			SNDDSKFGQRMLEKMGWSKGKGLGAQEQG
						ATDHIKVQVKNNHLGLGATINNEDNWIAHQ DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
			Į	İ		LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL
	}	ĺ				DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF
	}	}				TIOEYFAKRMAALKNKPOVPVPGSDISETOVE
						RKRGKKRNKEATGKDVESYLQPKAKRHTEG
		j]			KPERAEAQERVAKKKSAPAEEQLRGPCWDQ
						SSKASAQDAGDHVQPA
1328	2678	Α	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI
			l			CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF
1200	2670		10351		064	HMHCILKWLHAQQVQHCPMCRQEWKFKE
1329	2679	Α	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN
		1	}			LSFADICVTSTTIPKMLMNIQTQNKVITYIACL
			Ì		!	MQMYFFILFAGFENFLLSVMAYDRFVAICHP
		ļ	}			LHYMVIMNPHLCGLLVLASWTMSALYSLLQI
			ł			LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF
]			LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH
		1				AISSAQGKYKAFSTCASHLSVVSLFYGAILGV
ļ			ŀ		1	YLSSAATRNSHSSATASVMYTVVTPMLNPFI
1220	2600		10353	34	0672	YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680	Α	10352	34	2573	IPFLKSCCCCCLFDFPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESOSEDEOFENDLE
	i	1			1	TDPPNWOOLVSREVLLGLKPCEIKROEVINEL
		1	ļ	•		FYTERAHVRTLKVLDQVFYQRVSREGILSPSE
						LRKIFSNLEDILOLHIGLNEOMKAVRKRNETS
			1			VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ
			[PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR
						LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT
		1	i			EREKVKKAADHCRQILNYVNQAVKEAENKQ
1						RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK
İ		ĺ				RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV
						LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE
ĺ			1			LVAQTVSEKTVWQDLICRMAASVKEQSTKPI
						PLPOSTPGEGDNDEEDPSKLKEEOHGISVTGL
Í					•	QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD
[LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
						HLPVSEERWALDALRNLGLLKQLLVQQLGLT
[1	EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE
1						NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE
						SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
						MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
[!		ļ		GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP
						QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI
		ı				QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE
						SYTILCORLAGSALTDKHSDKS

CECUTAL CECUTAL MAY 1 CECUTAL TO CONTRACT OF THE CECUTAL CONTRACT OF THE CECUTAL CONTRACT OF THE CENTRACT OF THE CENTRAL CONTRACT OF THE CENTRAL CONTRACT OF THE CENTRACT OF THE CENTRAL CONTRACT OF THE CENTRAL CONTRACT OF THE CENTRAL CONTRACT OF THE CENTRAL CONTRACT OF THE CENTRACT OF THE CENTRAL CONTRACT OF THE CENTRAL CONTRACT OF THE CENTR	/A-A
	(A=Alanine C=Cysteine,
	Glutamic Acid, Glycine, H=Histidine,
eotide seq- USSN location corresponding l=Isoleucine, K=Lysi	
seq- uence 09/496 correspondi to last amino M=Methionine, N=A	
uence 914 ng to first acid residue Q=Glutamine, R=Arg	
amino acid of peptide T=Threonine, V=Val	
residue of sequence Y=Tyrosine, X=Unkr	nown, *=Stop codon,
peptide /=possible nucleotide	deletion, \=possible
sequence nucleotide insertion	
	PWPELGDAQPNPDKYLEG
	ETNKTDNTEAPVTKIELLP
	DDPWNLPTLQDSGIKWSE
	GRLILLLGFLYFFVCSLDIL
	GOFFSNSSIMSNPLLGLVIG
	SIVVSMVSSSLLTVRAAIP VALMQVGDRSEFRRAFA
	LVLLPVEVATHYLEIITOL
	PDLLKVITKPFTKLIVOLDK
	KNKSLVKIWCKTFTNKTO
	SLCWTDGIQNWTMKNVT
	IFHLPDLAVGTILLILSLLV
	VLKGQVATVIKKTINTDFP
	/GAGMTFIVQSSSVFTSAL
	PLTLGSNIGTTTTAILAAL
	LCHFFFNISGILLWYPIPFT
	SAKYRWFAVFYLIIFFFLIP
)	LVGVGVPVVFIIILVLCLR
	LQNWNFLPLWMRSLKPW
DAYVSKI IGCTQN	IRCCCCCRVCCRACCLLC DLEEAQEGQDVPVKAPET
FDNITISREAQGEV	
	SLLTAPHSLDLPALPPGPR
	SVKPSWGPGPSEGVTAVP
	DLFNHTLSECHVELSOST
1 1 1 1 1 1	VVGLVENLLVICVNWRG
SGRAGLMNLYILN	MAIADLGIVLSLPVWMLE
VTLDYTWLWGSFS	CRFTHYFYFVNMYSSIFF
	SASPSWQRYQHRVRRAM
	EVVHIQLVEGPEPMCLFM
	ALSTTILGFLLPFPLITVFN
1	KSRRHCLLLCAYVAVFV
1 1 1 1 1 1 1 1	TLHGTHISLHCHLVHLLY CVINPILYNFLSPHFRGRLL
1 : 1 1 1 1 1	AGTCASSSSCSTQHSIIIT
	EPSLSFQAHHLLPNTSPISP
TOPLTPS	
	ERASRAEPPAVAMGQND
	RVTKQYLPHVARLCLIST
FLEDGIRMWFQWS	EQRDYIDTTWNCGYLLA
	CVLVLSRNFVQYACFGLF
	LKFLMRNLALGGGLLLL
	GVPTMRESSPKQYMQLGG
	FDASFFSIVQNIVGTALMI
	LVVWLFAINVYFNAFW']
	OFFQTMSVIGGLLLVVAL
GPGGVSMDEKKKE 1334 2684 A 10367 59 1562 QAWSLQVALSPFFI	FPASPSNSFAAAVPQLLFP
	RRSARRFLIMSELTKELM
	OTIFCRWTQGFVFSESEGS
	PVQAFLLKKLLFSSEKSS
	CHTLCDILESACCDHSGS
1	ETASISGSPAESSCOVEHS
l l l l l l l l l l l l l l l l l l l	IALIQKRSFRSLPELKDAV
	VLLFLYSVLLTKGIENIKN
	SHGSQSLINLLLTGHAVSN
[[] [] [] [] [] [] [] [] [] [LLGIHEQAAVGFLTLMEA
LRYCKVGSYLKISK	IPYLDCLASETHLTVFFA
LRYCKVGSYLKISK KDMALVAPEAPSE	IPYLDCLASETHLTVFFA QARRVFQTYDPEDNGFIP VSDPEYINLMKNKLDPEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
1335	2685	A	10375	82	2929	DTPIKRCLQTKWPYIELLWTTDRSPSLN TRTKRRLGREKAMASPPRGWGGGELLLPFML LGTLCEPGSGQIRYSMPEELDKGSFYGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNANGKRDGKAPAGGNGN
1336	2686	A	10379	I	557	KKKSGKKEKK RPRRROPSFSCRVLVLEDPPCFRFTNSMNQEK LAKLQAQVRIGGKGTARRKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380		1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of NO: of nucl- eotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide seq- uence SEQ ID NO: of nucleotide sequence (A=Alanine of nucleotide sequence sequence sequence sequence (A=Alanine of nucleotide sequence (A=Alanine of nucleotide sequence	i, istidine,
nucle peptide eotide sequence uence	istidine,
eotide seq- uence USSN 09/496 corresponding uence uence 914 location corresponding to last amino mg to first amino acid residue of peptide residue of sequence y=Tyrosine, X=Unknown, *=Stop	•
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagine, P= quence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=Seria amino acid residue of peptide residue of sequence y=Tyrosine, X=Unknown, *=Stop	٠,
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=Serin amino acid residue of peptide residue of sequence Y=Tyrosine, X=Unknown, *=Stop	Proline.
amino acid residue of peptide residue of sequence residue of seque	
peptide / /=possible nucleotide deletion, \=po	
sequence nucleotide insertion	
QPN	
1339 2689 A 10386 50 390 LGAMAKHHPDLIFCRKQAGVA	IGRLCEKCDG
KCVICDSYVRPCTLVRICDECN	YGSYQGRCVI
CGGPGVSDAYYCKECTIQEKDI	RDGCPKIVNL
GSSKTDLFYERKKYGFKKR	
1340 2690 A 10388 113 3472 SQLRKGASATHSSPSRTDCIAQI	MMDIYVCLK
RPSWMVDNKRMRTASNFQWL	
NQVNSQKKGAPHDLKCVTNNI	QVWNCSWK
APSGTGRGTDYEVCIENRSRSC	YQLEKTSIKIP
ALSHGDYEITINSLHDFGSSTSK	•
SNVIWEIKVLRKESMELVKLVT	
TLHHWSWASDMPLECAHFVE	
GLEEWSDWSPVKNISWIPDSQT	•
VGSDITFCCVSQEKVLSALIGHT	
NVAIKIRNISVSASSGTNVVFTT	
AGYPPDTPQQLNCETHDLKEJIC LVGPRATSYTLVÉSFSGKYVRL	
YQLLFQMLPNQEIYNFTLNAHN	
VNITEKVYPHTPTSFKVKDINST	•
GNFAKINFLCEIEIKKSNSVQEQ	
NSSYLVALDKLNPYTLYTFRIRG	
SKWSNKKQHLTTEASPSKGPDT	
KNLIIYWKPLPINEANGKILSYN	
SLSEIPDPQHKAEIRLDKNDYIIS	•
SPPSKIASMEIPNDDLKIEQVVG	MGKGILLTW
HYDPNMTCDYVIKWCNSSRSE	PCLMDWRKV
PSNSTETVIESDEFRPGIRYNFFL	YGCRNQGY
	EDTSADSILV
KWEDIPVEELRGFLRGYLFYFG	
RVLESGRSDIKVKNITDISQKTL	•
YHLVLRAYTDGGVGPEKSMYV	
IIAILIPVAVAVIVGVVTSILCYR	
PDIPNPENCKALQFQKSVCEGS:	
CTPNNVEVLETRSAFPKIEDTEI	
RSDAKPENHVVESYCPPIIEEEIH TAQVIYIDVQSMYQPQAKPEEE	
GYKPOMHLPINSTVEDIAAEED	
QANVNTWNLVSPDSPRSIDSNS	
NSRQFLIPPKDEDSPKSNGGGW	
ND ND	
1341 2691 A 10392 1 5057 MLPPKHLSATKPKKSWAPNLY	ELDSDLTKEP
DVIIGEGPTDSEFFHQRFRNLIY	
IKLRNLCLDWLQPETRTKEENE	
PEKLKPWVRAKKPENCEKLVTI	
QPEGESLHGVLVVSAGLRCPLG	
SGLDNSLSWAAVGMSCVLWDI	
ATKSVSTHAQGDAAQGLGGTTV	/RMWARDSN
LATGVLLDDNNSDVTSDDDMT	
SVHSFSGDRDWDRRGRSRDTE	PRDRWSHTR
NPRSRMPPRDLSLPVVAKTSFEI	
RAYESRSQDAESYQNVVDLAEI	
NMENYRKLLSLGVQLAEDDGH	SHMTQGHSS
RSKRSAYPSTSRGLKTMPEAKK	
ESSHGVIMEKFIKDVSRSSKSGR	
RFPRMSDDNWKDISLNKRESVI	
FRGGFRFNSTLVSRKRVLERKR	
GSIHDQKGCPRKKPFECGSEMR	
SLSSPSFTESQPIDFGAMPYVCD	
EFVEHQIMHTRENLYEYGESFIL	ISVAVSEVQK

NO. of peptide sequence older by the peptide sequence of pepti	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Couling Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Sequence Dispays Contesponding Dispays Contesponding Dispays Contesponding Dispays Contesponding Dispays Contesponding Dispays			hod		, , ,	ľ	
uence 09/496 Orrespondi 1914 Ing to first			ĺ				
mino scid residue of peptide requence peptide sequence se		•	l				
residue of peptide sequence Y=Tyrosine, X=Ulxnown, **Stop codon, /*possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide delicin, *possible nucleotide insertion SVGGKKPECKDCGEFINSAALAEHRKHHA RGYLVECKNOGEEARMSPITSELQKIYGK DEFYECK VCESTETH, SALEHGKHIRGDDKD NERRHERERERERGETFRPSALNEFQKMYO NERRHERERERGETFRPSALNEFQKMYO NERRHERERERGESTFRPSALNEFQKMYO NERRHERERERGESTFRPSALNEFQKMYO NERRHERERGESTFRPSALNEFQKMYO NERRHERERGESTFRSSALSHOKHIRKIN LEGR GYESSVIHSIGSPTAMSPYON NERRHERERGESTFRSSALSHOKHIRKIN LEGR GYESSVIHSIGSPTAMSPYON NERRHERERGESSINISDLINDKROKIPAR HORVEAGGINSTGERSYSSIVISLUSKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGSGEFKKDGEFSVSSWINGKYFAKYOKSYPY GGGGSGEFKEFTGSTSSTHAPATATATATATATATATATATATATATATATATATAT	uence	j	1	914			
peptide							
	1	}				sequence	
SOVGGKPECKDCGETTH/SAALAEHRKIHA RGYLVECKNOCCEAR/PSPTSTELQKIYGK DKFYECRVCKETT-HSSALEH(OKIHGDDND NEREHRERERERGFFFFSPANLEFGKMYO KEKMYECKVCGETT-HSSSLEHOKHHRGDDND NEREHRERERERGFFFSPANLEFGKMYO KEKMYECKVCGETT-HSSSLEHOKHHRGDNDN PENKKGVCECTT-FIGOSLAKROKTYKIKELC DFTDGRDAPMOSSELSEHOKHHSRKNLEFGR GYEKSVHISGPTT-SSOKSHHTURPLESDEDEKA FTISSNPYENOKIPTKENVYEAKSYERSVHISL ASVERQKSHINAGPESDEDEKA FTISSNPYENOKIPTKENVYEAKSYERSVHISL ASVERQKSHINAGPESSYNSEVHSLVASEPPS HIGNELVESNEKGESSIVISLINDKROKIPAR ENPCEGGSKRINYEDSVIGSVFFAKPOKSYP GESGEFKKDGESSYSSIVISLIVAKSKPPS HORGLVESNEKGESSIVISLINDKROKIPAR ENPCEGGSKINSTHISLPGGOTFRRGMLYECO GEGEFKADGESTYSSSIVISLIVAKARAKK YÜBHRSINETSVHISLPGGOTFRRGMLYECO GEGEFKADGESTYSSSIVISLIVAKARAKK YÜBHRSINETSVHISLPGGOTFRRGMLYECO GEGEFKADGESTYSSSIVISLIVAKARAKK YÜBHRSINETSVHISLPGGOTFRRGMLYECO GEGEFKADGESTYSSSIVISLIVAKARAKK YÜBHRSINETSVHISLPGGOTFRRGMLYECO WEVIRSLAPTOPOTSVAOEQVAKPQARNIKCK WEVIRSLAPTOPOTSVAOEQVAKPQARNIKCK DFROFFATSEDLITINOKIYOQEKSTGEESGOE NTDIGETHISETHOQITEIPDVIGGSDMEDPO KOCCOGSTHISTVHITHSISEOVORDYTGEO UNFORCHSTENDEN KOCCODGITALL PMKPERNBAARERIPALLAGSA RICLLCOGGFHISSALMEHMAHREDDLLEON GWAAEEAHPGLAALTPEROPOTSERLFECAVCG SEFVINEALEAHDYTHKEEPPYVECKSTYTTS FLITEPLKCAPPYECKDOCKSTHISTVLTKHKE UNFORCHSTENDEN KOCCOGGFHISSALAAAAQEVEAHPULLEON GWAAEEAHPGLAALTPEROPOTSERLFECAVCG SEFVINEALEAHDYTHKEEPPYVECKAEPPE GRADEPPOCKOGESERHSTVLTKHKE UNFORCHSTENDEN FLITEPLKCAPPYECKDOCKSTHISTVLTKHKE LILEEEBEBBAAAAAAAAAAQEVEAHAFULLE GRADEPPOCKOGESERHSTVLTKHKE LILEEEBEBBAAAAAAAAAQEVEAHAFUL VARROGLANGEAAFEDRAAQOPORGAB GRADEPPOCKOGESERHSTVLTKHKE LILEEEBEBBAAAAAAAAAAQEVEAHAFUL VARROGLANGEAAFEDRAAQOPORGAB GRADEPPOCKOGESERHSTVLTKHKE LILEEEBEBBAAAAAAAAAAAAAAAAAAAAA GAABAAAAAA GAABAAAAAAAA				•			
RGYLVECKNÖEGERAFMF9FITSELQKIYOK DKFYECVCKETH-HEODDKD NEREHRERERERGETFR9F8ALNEFOKMYO NEREHRERERERGETFR9F8ALNEFOKMYO NEREHRERERERGETFR9F8ALNEFOKMYO NEREHRERERERGETFR9F8ALNEFOKMYO NEREHRERERERGETFR9F8ALNEFOKMYO NEREHRERERERGETFR9F8ALNEFOKMYO PENKOKVCKGETFIFGOSLKRROKTYNKERLC DFTDGRDAPMQSSELSEHOKHISRKINJEGR GVEKSVIHSGPTESQKSHTITRPLESDEDEKA FITSSPYPEROKLIPTKENYVPAKSYERSVIHSL ASVEAQKSHSVAGPSKPKVMAESTIOSEDANA HORVRAGGNTSGRBYSSSWHELVSKSPRSV HORDELVESKERGESSIYISDLINDKROKIPAR HORVELAGGNTSGRBYSSSWHELVSKASPRS HORDELVESKERDESSIYISDLINDKROKIPAR BENDEGGSKNRNYEDSVIQSVFRAKPOKYDA GROSGEFKKDGEFSYFSSNYREVGKARAKKK YIBHRSNETSVIHSLPFGGGTFRFROMLYECQ ECGCEAHSSDLIEHDKRIDERERGSSRNYE WSVIRSLAPTIDQTSYAQEGYAKEQARINKCK JPROFFATSEDLINTKKIVPQKESKIGEESGGE NTDGETHISETHIGQTTIEDPVIQGSDMEDPOQ KDDPDDKIYEGEDGGGFVDLITDLITDHQKVH SRKCLVDSREYTHSVHTHSISEVQRDYTGEQ LYECPKOGSFHISSTLEFHQRHEGDQLYSM KOCDDGFIALLPMKPRRNRAARENPALAGSA RICLLCGGGFHISSALHERHMRLHREDDLLEQS QMAEEAIIPGLALTFFQRSGTEERLFECAVCG ESFVNPRELADHTVYTKINEPPYECKDCASSFHISTVLTKHKE LHLEEEEEBDRAAAAAAAAQEVEANWIVPQ VVLRIQGLNVBAAEPSVBAABEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPVBABBEPBBABBABBABBABBABBABBABBABBABBABBABBABBA	-				Josquelle		
NRREHERERRERGETRPSSALNEFOKMYR (KEMPECKVOCETH-LISSIKERIQKHTRIGN FENKGKVCETTIPGQSI.KRRQKTYNKERU DFTOGNAFMOSES.SEHQKHISKIR.PGR GYEKSVHISGPTESOKSITTIRPLESDOESAK ATISSNYENOKUTYKERU PER GYEKSVHISGPTESOKSITTIRPLESDOESAK ATISSNYENOKUTYKERVYPRAKSYERSVHISU. ASVEAQKSHSVAGPEKPKVMAESTIQSEDAM HQRVRAGONTSEGREYSKSVHISU. VASKPPRS HNONELYESNEKGESYSTSUNIAL VASKPPRS HNONELYESNEKGESYSTSUNIAL VASKPPRS HNONELYESNEKGESYSTSUNIAL VASKPPRS GOGSGERKROPESVPSSWHRISU. VASKPPRS HNONELYESNEKGESSYTSDLINDKROKUPAM ENPCEGGSKRNYEDSVIQSVFRAKPQKAKK YTERRSNETSVHISULFOGGTFURROMLYECQUE EGGECFAHSSDLINDKROKUPAM ENPCEGGSKRNYEDSVIQSVFRAKPQKSWY GOGSGERKROPESVPSSWHRYOKARAKK YTERRSNETSVHISULFOGGTFURROMLYECQUE EGGECFAHSSDLINDKROKUPAM ENPCEGGSKERNYEDSVIQSVFRAKPOKAKK YTERRSNETSVHISULFOGGTFURROMLYECQUE EGGECFAHSSDLINDKROKUPAM ENPCEGGSHINDKROKUPAM ENPCEGGSHINDKROKUPAM ENPCEGGSHINDKROKUPAM ENPCEGGSHINDKROKUPAM ENPCEGGSHINDKROKUPAM ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGSHINDKROKUPAM ENGCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENGCAS ENPCEGGG ESPONDAL ENPCHANCE ENGCAS ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENPCEGGG ESPONDAL ENPCHANCE ENCOMENTAL ENCOMENTAL ENPCEGGG ENPCHANCE ENCOMENTAL ENPCEGGG ENPCHANCE ENCOMENTAL ENPCEGGG ENPCHANCE ENCOMENTAL ENPCE		ľ	ľ				
KEKMYECKVCGETELHSSILGEHQKIHTGRELC PFENKGKVCEETPIGOSLKRRQKTYNKERLC DFTDGRDAPMQSSELSEHQKIHSRKNILEGE GYEKSVHSGPTERGOSLKTRTPLESDEDEKA FTISSNPYENOKIPTKERVYPAKSYERSVHISLA ASVEAQKSHSVAPFERSVHSLTOSEDDAN HQRVRAGONTSEGREYSRSVHSLVASKPRS HMONELVESNEKGESSIYISDLNDRAQRIJPAG ENPECEGSKRRNYPDSVIQSVRTAKPÇKSVP GEGSGEFKKDGFFSVPSSNVREYQKARAKKY YIBHRSNETSVHBLIFGGTTEPRYGKAPÇKSVP GEGSGEFKKDGFFSVPSSNVREYQKARAKKY YIBHRSNETSVHBLIFGGTTEPRYGKAPKKY WSVRISLAPTDPQTSVAQEQYAKAPÇKSVP GEGSGEFKKDGFFSVPSSNVREYQKARAKKY TIBHRSNETSVHBLIFGGTTEPRYGQSSMEDPQ KDPDDKIYEGEDCGLGFVDLTDLTDHQKKY HORDETHSEETHGGTHEPRYGQSSMEDPQ KDPPDKYTSCEDCGLGFVDLTDLTDHQKSM KGCDDGFTALLEMFKRRNALAGGSA RICLLGGQGFHBSALNEHMRLHREDDLLGG QMAEALIFGLALTEPQRSTCEERLJECAVGG ESFVNPAELADHVTVHKNEPYEVGSSYTTHS FLEPLKGAPFYEKDCGKGSHFTSVLTKHKE LHLEEBEEDBAAAAAAAAPEVEAAPEV VALRIGGLINGBAAAPEVEAAPEVGAAPEV FLAGFPNGEAGPDGRAAFPGFAGOPNGBA QPNGGADEPDGAGEDPERAEPEGKAEPE GDADEPDGVGILDPEGGDDGIQVEFFYYDG HECTETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHLISLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHLISLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHTO HECHETFTSSTAFSEHLKTHASMIPPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCGQ LYNDHILSLARHQNTHASH PGSPEVFKLAHVTQQDEEHRCREGALSLLQ HAPPHPDPQEVLTLQFTDGTTTSLAGGARDPCATAGATAGA ASRRSPARPPVPAPPALFRGPGTEGSTSLS APAVLVVAVAVVVAVAVAVANAVHTIL PGSPEVFKLAHVTQQDEEHRCREGALSLLQ HAPPHDPQFEVLTLQFTDGTTSLAFFE EGGSTAFACHTORAGATAGATAGATAGATAGATAGATAGATAGATAGATAGA]						
PFENKGKVCETTIPGGSLKRRQKTYNKEKLC DFITOGRDAPMOSESLSEHQKURSKIPEGR GYEKSVHISGPTESQKSHTTIRPLESDEDEK AFTISSNYEMOKUTYKENVYEAKSYERSVHISL ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN HQRVRAGGNTSGGREVSKYERSVHISLVASKPPA HORNELVESNEKGESSIYISULNSKQKIPAR ENPCEGGSKRRNYEDSVIQSVFRAKPQKSVY GEOSGEFKKDGESTYSDLNDKRQKIPAR ENPCEGGSKRRNYEDSVIQSVFRAKPQKSVY GEOSGEFKKDGESTYSDLNDKRQKIPAR ENPCEGGSKRRNYEDSVIQSVFRAKPQKSVY GEOSGEFKKDGESTYSDNARPQKAKK YTEHRSNETSVHISLPFGEGTFRPRGMLYECQ EGGECFAHSSDLTEHQKHIDERSKPSGRNYE WSVIRSLAPTDPQTSVAQEQYAKEQARNKCK DPGQFTATSEDLNTNQKIYDQKSKIGEESQGE NTDGETHISETHGGETIEDPVQGSDMEDDQ KDDPDDKTYECEGCIGFVDLTDTHQKVH SRKCLVDSREYTHSVHITHISISEYQRDYTGGE LYCPCKGESFHISSTLEHQRHIERDQLLSSA RICLLGQGFHISSSLTEHQRHIERDQLLSSA RICLLGQGFHISSSLTEHQRHIERDQLLSSA RICLLGQGFHISSSLTEHQRHIERDQLLSSA RICLLGQGFHISSALNEHMAL HERDDLLGSA RICLLGGGFHISSALNEHMAL RICHDDAL HERDDLLGGFTTTSTTT THE THE THE THE THE THE THE THE THE							
DFTDGRDAPMQSSELSEHQUISRKNILFEGE GYEKSYHISGOPTESQKSHITTIPLESDEDEKA FTISSNPYENOKIPTKERVYPAKSYERSVIHSI ASVEAQKSHSVAPGKPEVPAKASTENSVIHSI ASVEAQKSHSVAPGKPEVPAKASTUGSEDAIN HQRVRAGGONTSEGREYSRSVIHSILVASKPPA HNGNELVESNEKGESSIYISID.NERQKIPAR ENPCEGGSKRRYPEDSVIQKRYKRAKPQKSVP GEGSGEFKKRGEFSVPSSNVREYQKARAKKY YIEHRSNETSVIHSILFTGEQTTEPRYGKAPAKKY YIEHRSNETSVIHSILFTGGTTEPRYGKIPAR WSVIRSLAPTDEQTISTPAQGTYARKGMLYECQ ECGECFAHSSDLTEHQKIHDREKFSGSRNYE WSVIRSLAPTDEQTSVAQEQYAKEQARNKCK DFRQFFATSEDLNINQKIVDQESSHGEESQG NTDGETHISEETHAGQTEIDEPVQGSSMEEDPQ KDDPPDKYYECEDCOLGFVDLTDLTDHQKYV SRKCLVDSREYTHAVIHTISISEYQADYTGEV LYCCPKCGSSFHSSTLFEHQRIHEQDQLYSM KCCDDGFTALLFMKPRRNRAAERRPALAGSA IRCLLCGQGFHISSALNEHMRI.HREDDLLGS QMAEARIJFGLALTEPQRSTCERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCCKSFHSTVLTKHKE LHLEEEEEDPAAAAAAAAQFVEANVHVOV VVLRIQGLINVEAAAPEVEAAAPEVEAAAPEV EAAPNGEAEGPOBAAAPIGEAQOPNOGA QPNGDADEPDGAGEDPERAEPEVEAAPEPV EAAPNGEAEGPOBAAAPIGEAQOPNOGA QPNGDADEPDGAGEDPERAEPEVGAAEPEV GDADEPDGVGIEDPERGEDDEIOVEFFYDDC HECTETTTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYPKCDVCOQ LINDHILSLARRQNTHTO HECTHETTSTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYPKCDVCOQ LINDHILSLARRQNTHTO LINDHILSLARRQNTHTO LINDHILSLARRQNTHTO VVLRIQGLINVEAAGAGADEPERAECPEGGAALSUL APAVLVVAVAVVVVSAVAVAMANNTHY PPGSPEVPKLINTVQDQEEHRCREGALSLU, QHAGGAPQLVCTIPGADEDINKRHLSIGFYVON THEDVVLVRIYGNKTELLVDRDEEVKSFRVL QHAGGAPQLVCTIPGADEDINKRHSJDFSSQI LQEEMTWMKELSNIGSSVVLCHINDLLCKNIN YNEKGODVGFIDVEYSGYNVLANSHFW GLWALIQAKVSTEFDFLGYAIVFRNQYFKM EKGCDVGFIDVEYSGYNVLANSHFW GLWALIQAKVSTEFDFLGYAIVFRNQYFKM KPETTALKVPE FAGYSDVDYSLYPDRELQSQWLRAYLEAYK GLWALIQAKVSTEFDFLGYAIVFRNQYFKM KPETTALKVPE LMYDTINGESNINAVQDWSTIGKTYSWDNAQ VILVGNKCOMEDERVISTERQHLEGQLOFG LMYDTINGESNINAVQDWSTIGKTYSWDNAQ VILVGNKCOMEDERVISTERQHLEGQLOFG FFETSKAKDNINVQCTFERLVDILICKMSSLEF FFETSKAKDNINVQCTFERLVDILICKMSSLEF FFETSKAKDNINVQCTFERLVDILICKMSSLEF FFETSKAKDNINVQCTFERLVDILICKMSSLEF FFETSKAKDNINVGCTFERLVDILICKMSSLIFF FFETSKAKDNINVGCTFERLVDILICKMSSLIFF FFETSKAKDNINVGCTFERLVDILICKMSSLIFF FFETSKAKDNINVGCTFERLVDILICKMSSLIF	1						
GYEKSVIHSGFTESOKSHTTRPLESDEDEKA FITSSNPYENGKIPTKENVYEAKSYERSVIHSL ASVEAGKSHSVAGPSKPKVMAESTIGSFAIN HQRVRAGGINTSGERSYRSVIHSLVASKPPRS HNGNELVESNEKGESSIYISDLNDKRQKIPAR ENPCEGGSKRINYEDSVIQSVFRAKPQKSVP GEGSGFKKDGEFSVPSSNVEQKARAKKK YIEHRSNETSVIHSLPFGEQITFRRGMLYECQ ECGECFAISSDLTEHQKHIDERKPSGSRNYE WSVIRSLAPTDPQTSVAQEQVAKEQARNKCK DFRQFFATSEDLNTNQKIYDQKSHGEESQGE NTDGEETHISEETHQGETIEDPVIQGSMBDPQ KDDPDDKIYECEDCGLGFVDLIDTDHQKVH SRKCLVDSREVTHSVHITHSISEYQRDYTGEQ LYFCKCGSSFIHSSFLFEHQKHIEQDOLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQGFHISSSALPEHMRYEGSSYTHTS FITEPLKGAIPPYECKDCGKSFIHSTVLTKHKE LHLEEEEDBAAAAAAAAQVEANVIPQ VVIRIQGINVEAAEPSVEAAEPEVGAAPFV EAAEPNGEAGOPDGEAAEPIGAAEPSVAGE GESGYIERASTSTGGANQADEKYPKCDVCQQ LYNDHLSLARRJNTHTG HERTERSTSFEHLKTHASMIPEPANAFG ECGGYIERASTSTGGANQADEKYPKCDVCQQ LYNDHLSLARRJNTHTG ASRSPAARPPVPAPPAPPERGGALSPLO GESGYVERASTSTGGANQADEKYPKCDVCQQ LYNDHLSLARRJNTHTD 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVGTRIGNSA ASRSPAARPVPAPPAPPAPPAPFRGGALSLQ HERTERTSFEHLKTHASMIPEPANAFG ECGGYIERASTSTGGANQADEKYPKCDVCQQ LYNDHLSLARRJNTHTG HERDVVLVRIYONKTELLVDRDEEVKSFRVL QARIGCAPQLYCTFRNGCLYEFIGGGALDFKH VCNPAIERLIARQLAKHAHAHANNOWIFN PPGSEPVENLIVYVQDQEEHRCGALSLQ HLRPHWDPQEVTLQLTFDGTTINKLIGCYVCN TMEDVVLVRIYONKTELLVDRDEEVKSFRVL QARIGCAPQLYCTFRNGCLYEFIGGGALDFKH VCNPAIERLIARQLAKHAHAHANNOWIFNL PRGQSDVQFIDYPSYSGYNTLAYDIGGNIFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLIPQNOWASA QULVANKYSTIEFDFLGVAIVRFNQYFKM KPEVTALKVPE GUWALIQAKYSTIEFDFLGVAIVRFNQYFKM KPEVTALKVPE GUWALIQAKYSTIEFDFLGVAIVRFNQYFKM KPEVTALKVPE GUWALIQAKYSTIEFDFLGVAIVRFNQYFKM KPEVTALKVPE EKRIKLQINDTAGGERYRTITTAYYRGAMGFI LMYDTINEESSROONPDMYRFLLIGNGSVGK TSFLFRYADDSFTSAFVSTVQIDFEKYKTVFKN EKRIKLQINDTAGGERYRTITTAYYRGAMGFI LMYDTINEESSROONPDMYRFLLIGNGSVGK TSFLFRYADDSFTSAFVSTVQIDFEKYKTVFKN EKRIKLQINDTAGGERYRTITTAYYROMAGFI LMYDTINEESSROONPDMYRFLLIGNGSVGK TSFLFRYADDSFTSAFVSTVQIDFEKYRTVFKN EKRIKLQINDTAGGERYRTITTAYYROMAGFI LMYDTINEESSROONPDMYRCHERDFSAGDGFI LMYDTINEESSROONPDMYRCHERDFSAGDGFI LMYDTINEESSROONPDMYRCHERDFSAGDGFI LMYDTINETSTORPPPQPNCAC TORMSTOR THE TORMSTOR TOR			[
ASVEAQKSHSVAGPSRFKYMAESTIGSPAIN HQRVRAGGINTSGREYSRSVHISLVASKPPRS HNGNELVESNEKGESSIYISDLADKRQKIPAR ENPCEGGKNRNYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGESSYISDLADKRQKIPAR ENPCEGGKNRNYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGESSYISDLADKRQKIPAR ENPCEGGKNRNYEDSVIQSVFRAKPQKSVP GEGSGEFKKDGESSVISDLADKRQKSVP GEGGGFKYBSSIVETYRYGKNYEQ ECGCFAHSSILTENQKIBERKPSGSRNYE WSVIRSLAPTDPQTSVAQEQVAKEQARNKCK DFRQFFATSEDLANTNQKIYDQKSHGEESQGE NTDGETHISEETHQETIEDPVIQGSMBDPQ KDDPDRYFEGEOGG.GFVALDTHORKYH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALIPMKFRRNRAARERNPALAGSA IRCLLCQGFIHSSALNEHMRLHREDDLLEQS QMAEAHIPGLALTEP(RSQTER)FECAVCG ESSVNPABLADHVTVHKNEPYSYGSYVTHTS FLITEPLKGAUPFYECKDCGSFISTVLTKHKE LHLEEEEDRAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE LHLEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE GEGSVJERASTSTGGANQADEVAAPISVLASHEPPV AAAPISCALASHEPVEAAPISTVLTKHKE FLITEPLKGAUPFYECKDCOGSFINATVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHVPQ VVLRIQGI.NVEAAPEPVEAAPISTVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHPQ VVLRIQGI.NVEAAPPVEAPALAPEVAAPISTVLTKHKE CHLEEEEDRAAAAAAAAQEVEANVHPQ VVLRIQGI.NVEAAPPVEAPALAPGEAAPISTVLTKHKE CHLEEEDRAAAAAAAAQEVEANVHPQ VVLRIQGI.NVEAAPPVEAPALAPEGAAPISTVLTKHKE CHLEEEDRAAAAAAAAAQEVEANAPISTVL CHLEEEDRAAAAAAAAQEVEANAPISTV 1342 2692 A 10393 2 1350 GRRSSSDINFMFLERGAGI.SSAAVQTTRIGNSA ASRRSPAARPPVPAPPALPRGRGTGGSTSLS. APAVLVAVAVVVVVAWAMANNYHP PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HR.PHWDPQEVTLQ.FTDGITNKLIGCVYQN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL. QAHGCAPQL.YCTFNNGLCYEFIQGEALDPH VCPAIRFLIRAPLASHIPNICHTGGTADEDNKRTLSDPSQI LQEEMTWMKEUDJKVJPPBELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETGFLGYATVRNQYFKM KPEVTALKVPU LQEAGCAPQLYCTTRNQGEPKTITTVRQTYFKM KPEVTALKVPU EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEETGH.GYVLCHNDLLCKNIII VPKQGGDQFDYSYLJCHDDLCKMSSLGT EFKGSAKDINNVKQTFERLGDHLIGKMSSVGK TSFLFRYADDSFTSAFVSTVGIDFKKYTVFKN EKRIKLOJWDTAG			1				
HQRVRAGGNTSEGREYSRSVHISLVASKPPRS HNGNELVESNEKGESSIYSBLNDKRQKIPAR ENPCEGGSKNRNYEDSVIQSVFRAKPQKSVP GEGSGFKKDGEFSVPSSNVREYQKARAKKK YIEHRSNETSVHSLPFGEGTIFRFGMLYECQ ECGECFAHSSDLTEHQKHIDEKFSGSRNYE WSVIRSLAPIPOTISVAGEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGETHISEFTHGQTTEIDPVIQGSDMEDPQ KDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHHTSIEQFOLYSM KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRRNARAFPALAGSA KGCDGFALIPMFRFRALIPMFRFRALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KGCDGFALIPMFRFRAMAFPALAGSA KASASASAANAFALIPMFRAMAFPALIPMFRAMAFPALIPMFRAMAFPALIPMFRAMAFA KASASASASAANAFALIPMFRAMAFPALIPMFRAMAFA KASASASASASANAFATAFPALAGAMAFA KASASASASANAFATAFPATAFA KASASASANAFATAFPATAFALIPMFRAMAFA KASASASASANAFATAFATAGAMAFA KASASASASAANAFATAFATAFATAGAMAFA KASASASAANAFATAFATAGAMAFA KASASASAANAFATAFATAGAMAFA KASASASAANAFATAFATAGAMAF	1					}	FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
HNÖRELVESNEKGESSIYISDLNDKRQKIPAR ENPCEGGSKRNYNDENVYFRAKPQKSVP GEGSGEFKKDGEFSVPSSNVREYQKARAKKK YIEHRSNETSVIRSLPFGEQTFRRRGMLYECQ ECGECHSSDLTEHQKHEREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLTINQKIYQEKSHGEESQGE NTDGEETHSEETHGGETIEDPVIQGSMEDPQ KDDPDBKIYDCEDCGLGFUDLTDLTDHGKVH SRICLVDSREYTHSVIHTHSISEYQRDVTGEO LYECPKCGESFHSSTLEHGBLIECDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQGFHBSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEPGRSGSTHIST FLTEPLKGAJRPYECKDCOKSFHISTVLTKHKE LHLEEEEDBAAAAAAAAQEVEANTHVPQ VVIRIGGINVEAAFEVEAAFPEVEAAEPEV EAAEPNGEAEGPDGBAAEPIGAGQPNGEAB QPPNGDADEPDGGGIEDPEGAGQPNGEAB QPNGDADEPDGGGIEDPEGAPEVEAAEPEV EAAEPNGEAEGPDGGAGEPPGFAAEPE GDADEPDGVGIEDPEGGEDQEIQVEEPYYDC HECTETTTSTSTAFSEHLKTHASMIEPFANNFG ECSTYLBRASTSTGGANQADEKYFKCDVCOQ LYNDHLSLARHQNTHTG GPRSSSDNNFINLERFAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS ASRRSPAARPPVPAPPALPRGRPGTGSTSLS APALVVAVAVVVVVVSAVAWAMANYIEV PPGSFEVFKLNYTVQDQEBERCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVREDEEVKSFRVL QAHGCAPQLYCTFNNGLCYETGGALDFKH VCNPAIFFILIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTIGFADEDINKRFLSDIPSSQI LQEMTWMKELISNLGSWU,RAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKGRGFTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKGRGFTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKGRGFTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKGRGFTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM EFKRIKLQIPUDTAGGERYRITAYYRGAMGFI LMYDITNEESPNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFGLGFG FFETSAKDNINVRQTFFERLTITAYYRGAMGFI LMYDITNEESPNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGGLGFG FFETSAKDNINVRQTFFERLTILLESKGHISPAKDTS 1344 2694 A 10395 2 4136 DRPPWNSRVDDPVTILLIELSKGHISPAKDTS							l . •
BNPCEGGSKNRNYEDSVIQSVFRAKKKÖKSVP GEGSGGFEKDGFESVPSSNYETYOKARAKKK YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ ECGECHASSOLTEHQKIHDREKPSOSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSKGESQGE NTDGETHESETHIQGETIEDPVIQSDMEDPQ KDPPDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVIHTSEYQRDYTGFQ LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMRFRRNRAAERPALAGSA IRCLLCGGGFIHSSALNEMRIHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFCAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIFFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEPEV EAAEPMGEAEGPDGEAAAPICEAACPOWEAAE QPNGDADEPDGAGIEDPEERAEEPEGKAEPEV GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETTTSSTAFSEHLKTHASMIFEPANAFG ECSGYUERASTSTGGANQADEKYFKCDVCGQ LFNDHISLARHQNTHITG ECSGYUERASTSTGGANQADEKYFKCDVCGQ LFNDHISLARHQNTHITG HECTETTTSSTAFSEHLKTHASMIFEPANAFG ECSGYUERASTSTGANQAAMAANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGSTSLS APARLVVAVVVVVVVSAVAWAMANYIHV PROSPEVPKLINTVQDQEEHRCREGOSTLSUP PROGRAMAE ECSGYUERASTSTORDHAMATALETTPROPORACA ULVANACAMA							
GEGSGEFKKDGEFSVPSSNVREYQKAÄAKKK YIEHRSNETSVIHSLPFGETPRRGMLYECQ ECGECPAHSSDLTEHQKIHDREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DPROFFATSEDLTINQKIYDQEKSHGEESQGE NTDGEETHSEETHGGETIEDPVIQGSMEDPQ KDDPDBKTVECEDCGLGFULTDLTDHOKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNAAERNPALAGSA RICLLCGQGFIHSSALNEHMRIHREDDLLEQS QMAEEAIIPGLALTEPGRSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEPEVAAEPEVAAEPEVAAEPEV EAAEPNGEAAGPDGEAAEPEGAQPNGGAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEEGGEAQQPNGGAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEEGGEAQQPNGGAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEEGGEAQQPNGGAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEEGGEAQPNGEAE APAVLVAVAVVVVVSAVAWAMANYHHV PROSPEVPKLINTVODQEBEHCREGAGISSAAVQTRIGNSA ASRSFAARPPVAPPALPRAPRGRPGTEGSTSLS APAVLVAVAVVVVVVSAVAWAMANYHHV PPGSPEVPKLINTVODQEBEHCREGGALSLLQ HLRPHWDPQEVTLQLFTDGITINKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSRFVL QAHGCAPQLYCTFINNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHAHANGHYENSIL WKMGKYSILDFIGABDDINKRISDIPSSQI LQEEMTWMKEILSULGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIGVNGFALASHFFW GLWALIQAKYSTIEFDFIGADGVAINFNSDESQU LQEEMTWMKEILSULGSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIGVNGFALASHFFW GLWALIQAKYSTIEFDFIGADGVAINFNSDHYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QUANYGGKDSSDQNFTYMFKLLIGNSSVGK TSILFRYADDSFTSAFVSTVGIDFKKVTVEKN KPEVTALKVPE EKRIKLQIWDTAGGGERTRETTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCOMEDERVISTERGTUTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCOMEDERVISTERGTUTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCOMEDERVISTERGUHLESKGFULDGLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLEKGFIIFSPAKDTS							
YIEHRSNETSVIRSLPFGEQTFRPRGML/YECQ BCGECHAPSDLTEHQKHDEREPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLTNIQKIYDQEKSHGESQGE MTDGETHESETHQGETIEDPVIQGSDMEDPQ KDPDDKIYECEDCGLGFVDLTDLTDHGKVH SRKCLVDSREYTHSVIHTHSEYQRDYTGEQ LYECPKCGESFIHSSTLFEHQRIHEQDQLYSM KGCDDGFIALIPMKPRRAAERNPALAGSA IRCLLCQGFIHSSALNEHMRLHREDDLLEQS QMAEBAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAJPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIGGLNVEAAFEVVAAAPEVGAAAFEV EAAEPNGEAAGPOPGAAFECKOCKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIGGLNVEAAFEVGAAFEVGAAEPEV GAAEPNGEAAGPOPGAAFECKAGPPGGEA QPNGDADEPDGGIEDPEGGEODGIQVEEPYYDC HECTETTISTSTATSEHLKTHASMIIFEPANAFG ECSGYURRASTSTGGANQADEKYFKCDVCQQ LFNDHLSLARHQNTHIUF PGSFEVFKLARYVYQVVVVSAVAWAMANYIHV PGSFEVFKLANYVYQDQEBHRCREGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYHVY PGSFEVFKLNVTYQDQEBHRCREGALSLLQ HLRPHVDPGVTUQLTTDGTINKLIGCYYGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVFGQGALDPKH VCNPAIFELIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDLGNHFNE FAGVSSVDVSLYPDRELQSWURAYLEAYK EFKGFGTEVTEKEVBILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FRGVSSVDVSLYPDRELQSWURAYLEAYK EFKGFGTEVTEKEVBILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FAGVSSVDVSLYPDRELQSWURAYLEAYK EFKGFGTEVTEKEVBILFIQVNQFALASHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FRGRGTEVTEKEVBILFIQVNQFALASHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FAGVSSVDVSLYPDRELQSWURAYLEAYK EFKGFGTEVTEKEVBILFIQVNQFALASHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FRGRGTEVTEKEVBILFIQVNQFALASHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FRGRGTEVTEKEVBILFIQVNQFALASHFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FFGFGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKEVBILFIQVNGFALSHFW FRGRGTEVTEKTEVBILFIQVNGFALSHFW FRGRGTEVTETTON FRGRGTT FRGRGTTTTTTTTTTTTTTTTTTTTTT							
WSVIRSLAPIDPQITSYAQEQYAKEQARNIKCK DFROFFATSEDLINTNQKIYDQEKSHGESQGE NTDGETHSETHIQETHEDPVIQGSSMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKYH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALI PMKPRNAAERNPALAGSA IRCLCQQGFIHSSALNEHMRLHREDDLLEQS QMAEEAIIRGLALTEFQRSQTEERLFECAVCG ESFVIPPAELADHVTVHKNEPYEYGSSYTHTS FLITEPLK.GAIPPYECKDGGKSFIHSTVLTKHKE LHLEEEEDBAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAA QPNGDADEPDGAGIEDERAEPERGKAEEPE GDADEPDGVGIEDPEEGEDDGIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIFEPANAFG ECSGYLERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ECSGYLERAFSTSTGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG SRRSPAARPYPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNYTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVYLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFFRLIARQLAKIHAHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEXQGDVQFIDVSYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVYTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTEETPGLGYAIVRNQYFKM KPEVTALKVPE EKRIKLQIVDTAGQERYRIITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNCOMEDERYNTGOPHACK TSFLFRYADDSFTSAFVSTVGIPFKVKTVFKN EKRIKLQIVDTAGQERYRIITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNCOMEDERYSTERGQHLGFEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPQPNCAC DRPPWNSRVDDFYVTILHLSSKGHISFAKDTIS							-
DFROFFATSEDINTNOKIYDQEKSHGESQGE NTDGETHSSETHOGHTEDPVIQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHISISEYQRDYTGEQ LYECFKGESSHIASFLERQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRARARPALAGSA IRCLLGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNPALAGSA IRCLLGGGGFHISSALNEHMERNHERPGGSSYHHTS FLIFELKGAPPYECKGSSYHHTSI FLIFELKGAPPYECKGSSYHHTSI FLIFELKGAPPYECKGAPEPVE GAAEPNGEAGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEGAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEGAAEPIGEAGQPNGEAE GCSGYIERASTSTGGANQADEKYPKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSFAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYHIV PRGSPEVPKLNVTVQDGEPKRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYYGG TMEDVVLVRIVGNKTLLVPRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGALDPKH VCNPAIFRLIARQLAKHHAHARINGWIPKSNL WLKMGKYFSLPTOFADEDINKFFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHDLLCKNII VNEKQGDVOFIDVSFSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTLETPFLGYAIVFRNOYFKM KPEVTALKVPE FAGTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLIFRYADDSFTSAFVSTVGIDFVKTVVFKN KPEVTALKVPE EKRIKLQIVDTAGGGERGYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTGIKTYSWDNAQ VILVGNKOMBEDERVSTERGGOLGFG LGFE FFETSAKDNINVKQTFFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPQNACC DPAPMNSRVDDFVTNLHLSSKGHISPAKDTS	1						•
NTDGETHISETHGOÉTIEDPÜQGSDMEDPQ KDDPDKIYECEDGGFVDLTDLTDHGKVH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESFIHSSTLEFHGHEQDQLYSM KGCDDGFIALLPMKPRRNRAGERNPALAGSA IRCLLCGGGFHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPABLADHVTHKNEPYEYGSSYTHTS FLIEPLKGAIPFVECKDGGKSFHSTVLTKHKE LHLEEBEDEAAAAAAAAQEVEAANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVE EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIDDPEEGEDDEQIQVEEPYYDC HECTETTSTSTAFSEHLKTHASMIIFPEANAFG ECSGYIBRASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ECSGYIBRASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRBRFGTEGSTSLS APAVLVVAVAVVVVSVAVAWAMANYHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPGEVTLQLFTDGTINKLIGCVVGN TMEDVVLRYTGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVFFIQGEALDPKH VCNPAJFRLIARQLAKHAHHAHNGWPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKGGDVQFIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYEYSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLCHNDLLCKNII YNEKGGDVGPIDYETSGYNVLGNGNSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGGRYRITITAYYRGAMGFI LMYDITNEESSNAVQDWSTOJGKTYSWDNAQ VILVGNKCDMEDBESNATGGOPLAGGGLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVNLHILLSSKGHISPAKDTS							
SRDCPDDKIYECEDCGLGFVDLTDLTDHQKYH SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ LYECPKCGESPHISSLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCGGFHISSLANEHMRLHREDDLLCGSA IRCLLCGGFHISSLANEHMRLHREDDLLCGSA IRCLLCGGFHISSLANEHMRLHREDDLLCGSA IRCLLCGGFHISSLANEHMRLHREDDLLCGSA GESFVNPAELADHVTVHKNEPYEVGSSYTHTS FLTEPILKGAIPPYECKDCGKSFHISTVLTKHKE LHLEEBEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPVEAAEPE							
SRKCL/DSREYTHSVHITHSISEYQRDYTGEQ LYECPKCGESFIHSSTLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQGFIHSSALNEHMRHREDDLLEQS QMAEAHIPGLALTEFQRSQTEERLFECAVCG ESFVNPABLADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAEFEVEAAEFEVE EAAFPRGEAGEPDGAAAFICASGQPNGEAE QPNGDADEPDGAGIEDPEERGEDQEIQVEEPYYDC HECTEITFTSSTATSFEHLKTHASMIHFEPANAFG ECSGYIBRASTSTGGANQADEKYFKCDVCCQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRFGTEGSTSLS APAVLVVAVAVVVVSVAVAMAMANYHVY PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDNKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYSGYNYLAYDIGNHFNE FAGVSDVDVSLYPDRELQSQWLRAYLEAYK EFKGFTEVTEKEVELIFQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDWFKLLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGGERYRITITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDBERVISTERGGHLGGLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPQNCAC							, , , , , , , , , , , , , , , , , , , ,
KGCDDGFIALLPMKPRRNRAAGRNPALAGSA IRCLLCQQGFIHSSALNEHMRLHREDDLLEQS QMAEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCOKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAACPEVEAACPEVEAACPEVEAACPEVEAACPOW EAAEPNGEAEGPDGEGAEGPGEAGQPNGEAE QPNGDADEPDGAGIEDPEEGEAGEPEGBAAPEGBAEPEGBA							
IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAPPYECKDCOKSFIHSTVLTKHKE LHLEEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETTSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERGGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDFKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFFM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDREVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICCKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVYNLHILBLSSKGHISPAKDTS							
QMAEEAIPGLALTEPQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FI.TEPLKGAIPYECKDCGKSFIHSTVLTKHKE LHLEEBEEDBAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAEPEVEAAEPEVEAAEPEVE EAAEPNGEAGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGVGEDPEGEDQEIQVEEPYYDC HECTEITISSTAFSEHLKTHASMIIFEPANAFG ECSGYIBRASTSTIGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSFAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAMAMNYHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITINKLIGCYVGN TMEDVVLVBYIGNKTELLVDRDEEVKSFBVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFJDTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYETYSGYNYLAYDIGNHFNB FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVFRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNITRLKETPPPPPCPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLISSKGHISPAKDTS							l l
ESFVNPAELADHVTYHKNEPYEYGSSYTHTS FITEPLKGAIPFYECKDCOKSFIHSTVLTKHKE LHLEEEEEDBEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVBAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAGEPEGAGEPE GDADEPDGVGIEDPEEGEDGEJQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSDNRNFLRERGLSSAAVQTRIGNSA ASRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVAVAVVVVVSAVAWAMANYHV PPGSPEVPKLNVTVQDQEHRCREGALSLLQ HLRPHWDPGEVTLQLFTDGTINKLIGGVVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEEIQGEALDPKH VCNPAIFELIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQTJSYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFQVNQFALASHFFW GLWALIQAKYSTTEFDFLGYAIVFRNQMSAS AGARYGAGROSDOMFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLIPTMRHEAPMQMASA QDARYGGKDSDOMFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN ERRIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPOPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLISSKGHISPAKDTS							
FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGIAVEAAAFPEVEAAFPEEAFPEGAEAFPE GDADEPDGAGIEDPEEGEDQEIQVEEPYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHITG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPOGSPEVPKLNNTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGINKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYFFIQGALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMELISNLIGSPVVLCHHDLLCKNII YNEKQGDVGIDYEYSGYVVLCHNDLLCKNII YNEKQGDVYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGQMSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRKLQIWDTAGQERYRITITTAYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQDFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
VVLRIQGLNVEA.AEPEVEA.AEPEVEA.AEPEV EAAEPNGEAEGPDGEA.AEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGAGIEDPEEGEDQEIQVEEPYYDC HECTETITISTATAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLEERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVRDREEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHAHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTYMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHENE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVXQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPPEGKAEEPE GDADEPDGYGIEDPEGEDQEIQVEEPYYDC HECTETITSTATSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLEERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERQQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							,
QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGYOGIEDPEGEDQEIQVEEPYYDC HECTETTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYESYSGYNYLAYDIGNHENE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPOPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSKGHISPAKDTS			!				,
GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETTSTSTAFSEHLKTHASMIIFERANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAHFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKGGDVQFIDYEYSGYNYLAYDIGNHENE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSKGHISPAKDTS							
HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNITHTG 1342 2692 A 10393 2 1350 GRPSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1						, -
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAMAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFIDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPONCAC							` ` '
ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPPQNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	12.40	0.000		10000			
APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVIT.QLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITTTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1342	2692	A	10393	2	1350	,
PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC	1						
HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNIII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC							
QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNIHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNIII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS						•	
LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							•
YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1					,	
1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							,
QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1343	2693	A	10394	102	839	
TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	"		''	10074			
LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							
VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS		ļ					
FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS		l					
DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							7
1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS						į	
	1344	2694	Α	10395	2	4136	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	sequence	/=possible nucleotide deletion, \=possible
1 .				sequence		nucleotide insertion
				Sequence		RKLOGKLPELOGVETELCYNVNWTAEALPSA
1]	į	
					Ì	EETKKLMWLFGCPLLLDDVARESWLLPGSN
1						DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV
						DRVETTRRYRLSFAHPPSAEVEAIALATLHDR
1						MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR
						LALEKANQELGLALDSWDLDFYTKRFQELQR
						NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG
						QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA
1 1						IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVFT
						AETHNFPTGVCPFSGATTGTGGRIRDVQCTG
1 1						RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF
1						QYPGNFARPLEVAIEASNGASDYGNKFGEPV
						LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS
] .						MEADHISKEAPEPGMEVVKVGGPVYRIGVGG
						GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ
						KMNRVIRACVEAPKGNPICSLHDQGAGGNG
						NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW
1 1						GAEYQESNALLLRSPNRDFLTHVSARERCPA
1 1						CFVGTITGDRRIVLVDDRECPVRRNGQGDAP
1						PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP
1						MLQPLALPPGLSVHQALERVLRLPAVASKRY
[[LTNKVDRSVGGLVAQQQCVGPLQTPLADVA
)						VVALSHEELIGAATALGEQPVKSLLDPKVAA
1 1						RLAVAEALTNLVFALVTDLRDVKCSGNWM
1						WAAKLPGEGAALADACEAMVAVMAALGVA
1						VDGGKDSLSMAARVGTETVRAPGSLVISAYA
1 1						VCPDITATVTPDLKHPEGRGHLLYVALSPGQ
						HRLGGTALAQCFSQLGEHPPDLDLPENLVRA
1 1		' i				FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM
1			:			AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV
1 1		1				LEVQEPDLAQVLKRYRDAGLHCLELGHTGE
1	1					AGPHAMVRVSVNGAVVLEEPVGELRALWEE
1 1		- 1				TSFQLDRLQAEPRCVAEEERGLRERMGPSYC
1 1	j	}	.]			LPPTFPKASVPREPGGPSPRVAILREEGSNGDR
1 1						EMADAFHLAGFEVWDVTMQDLCSGAIGLDT
]]						FRGVAFVGGFSYADVLGSAKGWAAAVTFHP
1 1						RAGAELRRFRKRPDTFSLGVCNGCQLLALLG
]		ļ				WVGGDPNEDAAEMGPDSQPARPGLLRHNL
]		l				SGRYESRWASVRVGPGPALMLRGMEGAVLP
1	İ	l	i			VWSAHGEGYVAFSSPELQAQIEARGLAPLHW
						ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR
	1	ı		l		HLAVMPHPERAVRPWQWAWRPPPFDTLTTS
						PWLQLFINARNWTLEGSC
1345	2695	A	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRE
	2073	^	10570	~ <i>_</i>	U72	RVAMHYOMSVTLKYEIKKLIYVHLVIWLLLV
1	1	ŀ				
]))				AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI
•]	j				LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI
	j	}				YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI
] [1					MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
1.246	2606		10000			QEKKHK
1346	2696	Α	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS
į 1		ł		Ì	Ì	GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT
			ì	ł		EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL
	Ì	1		I		TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL
	}	ŀ		J	ľ	NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK
		1				TVGIDDLTGEPLIQREDDKPETVIKRLKAYED
	Į	ļ	ì		ļ	QTKPVLEYYQKKGVLETFSGTETNKIWPYVY
		l				AFLQTKVPQRSQKASVTP
1347	2697	A	10402	153	1969	KHRQENNALDMAPEIHMTGPMCLIENTNGEL
	[VANPEALKILSAITQPVVVVAIVGLYRTGKSY
						

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		{		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ	[residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	!			peptide	35423355	/=possible nucleotide deletion, \=possible
1		ł		sequence		nucleotide insertion
						LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
1						PHPKKPEHTLVLLDTEGLGDVKKGDNONDS
	ĺ	Ì				WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY
	}					VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
ı	l	ł	1			WTLRDFSLDLEADGOPLTPDEYLEYSLKLTO
į						GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
1			}			HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
i	1					FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
1	ļ					GDLPCMENAVLALAQIENSAAVQKAIAHYD
	[QQMGQKVQLPAETLQELLDLHRVSEREATEV
}						YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
				•		QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
1	}					PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
	Į.					OTYLKSKESVTDAILOTDOILTEKEKEIEVEC
1	}					VKAESAQASAKMVEEMQIKYQQMMEEKEKS
i	ŀ					YOEHVKOLTEKMERERAOLLEEOEKTLTSKL
1	}	ł				QEQARVLKERCQGESTQLQNEIQKLQKTLKK
1		1				KTKRYMSHKLKI
1348	2698	Ā	10404	5	892	TOLPAPLSGVLSRLOLGSGAPLLTWVOETAG
1340	2070	*	10.01	,	0,2	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
1						DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
}						RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
1	ļ					EGNFAILALGGGYLHWGHFEMMRLTINRSM
1	}					DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
	1					GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
1	ĺ					FLDOVAHKLPFAAKAVSRGTLEKMRKDOEE
						RERNNONPWTFERIATANMLGIRKVLSPYDL
	1					THKGKYWGKFYMPKRV
1349	2699	A	10409	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
		i i				AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
						NTRGGTREIGSALTRMCMRHRSIEAKLROFSS
						ALIDCLINPLOEQMEEWKKVANOLDKDHAK
						EYKKAROEIKKKSSDTLKLOKKAKKGRGDIO
						POLDSALODVNDKYLLLEETEKOAVRKALIE
						ERGRECTEISMLRPVIEEEISMLGEITHLOTISE
1	[·			DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
	l					YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
						GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
				,		DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
						PDPNGGGPTTASGPPAAAEEAQRPRSM
1350	2700	A	10410	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
1.200] _,,,,					RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
						AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
1	!					
	•					K T I T F A L K C.P.A F K P M V C.SFF U K (VII JK INK I JI V V
						KYTYPALREEAPREHVESFFQKMDRNKDGV VTIEEFIESCQKDENIMRSMQLFDNVI

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

- 13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.